

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2002, 04:57:01 : Search time 1875 Seconds

(without alignments)
5937.556 Million cell updates/sec

Title: AI327498
Perfect score: 532
Sequence: 1 AATTCGCACTAGACACCC.....CATTAATTCCTGTCATCAT 532

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_hgtgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
------------	-------------	-------	--------	-------	-------------

1	511.6	96.2	1412	10	AF079835	AF079835 Mus muscu
2	511.6	96.2	1475	10	AF079834	AF079834 Mus muscu
3	325.2	61.1	175239	2	AL606903	AL606903 Mus muscu
4	325.2	61.1	226957	2	AL606971	AL606971 Mus muscu
5	101.8	19.1	1316	9	HMWCA6	HMWCA6
6	98.6	18.5	1344	4	BRCARANVI	BRCARANVI
7	68.6	12.9	107714	2	AL356306	AL356306 Homo sapi
8	68.6	12.9	141428	9	AF128411S7	AF128411S7 Homo sapi
9	67	12.6	432	9	AF128411S7	AF128411S7 Homo sapi
10	42.4	8.0	188036	2	AC091984	AC091984 Homo sapi
11	42.4	8.0	195909	9	AF165124	AF165124 Homo sapi
12	42.4	8.0	209836	9	AC091926	AC091926 Homo sapi
13	41.6	7.8	14568	1	AE004453	AE004453 Pseudomon
14	41.4	7.8	1077	9	AF128411S6	AF128411S6 Homo sapi
15	40.6	7.6	82576	2	AC098147	AC098147 Homo sapi
16	40	7.5	28142	9	HSL58B6	HSL58B6
17	39.4	7.4	157665	10	AC069018	AC069018 Mus muscu
18	39.2	7.4	75314	2	AC023378	AC023378 Homo sapi
19	39	7.3	142376	2	AC016959	AC016959 Homo sapi
20	39	7.3	178733	2	AF238577	AF238577 Homo sapi
21	38.8	7.3	5145	2	AC015315	AC015315 Drosophi
22	38.8	7.3	5284	3	DDOKRAGEN	DDOKRAGEN
23	38.8	7.3	7218	6	AC012373	AC012373 Drosophi
24	38.8	7.3	182080	3	AE003509	AE003509 Drosophi
25	38.8	7.3	303760	3	AE003509	AE003509 Drosophi
26	38.6	7.3	91822	9	AL160268	AL160268 Homo sapi
27	38.6	7.3	134499	2	AC096084	AC096084 Rattus no
28	38.6	7.3	160408	2	AC012196	AC012196 Homo sapi
29	38.4	7.2	196922	2	AC098158	AC098158 Rattus no
30	38.2	7.2	211301	2	AC100752	AC100752 Mus muscu
31	38	7.1	2426	2	AC020448	AC020448 Drosophi
32	38	7.1	169649	9	AL133328	AL133328 Drosophi
33	38	7.1	172904	3	AC007414	AC007414 Human DNA
34	38	7.1	260910	3	AE003831	AE003831 Drosophi
35	37.6	7.1	148415	9	AC016670	AC016670 Homo sapi
36	37.6	7.1	149572	2	AC104780	AC104780 Homo sapi
37	37.6	7.1	149768	2	AC016239	AC016239 Homo sapi
38	37.6	7.1	184831	2	AC016030	AC016030 Homo sapi
39	37.4	7.0	4969	8	AB017039	AB017039 Homo sapi
40	37.4	7.0	7606	8	AB034966	AB034966 Schizosac
41	37.4	7.0	7743	8	SPAP9598	SPAP9598 Schizosac
42	37.4	7.0	69774	2	AC105327	AC105327 S. pombe
43	37	7.0	2898	8	AF170173	AF170173 Mus muscu
44	37	7.0	5101	1	PSEMLABC	PSEMLABC
45	37	7.0	23630	1	AE008803	AE008803 Salmone

ALIGNMENTS

RESULT 1
AF079835
LOCUS
DEFINITION Mus musculus secreted carbonic anhydrase isozyme VI precursor.
ACCESSION AF079835
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1412)
REFERENCE
AUTHORS Wang, X.Z., Kuroda, M., Sok, J., Batcharova, N., Kimmel, R., Chung, P.,
Zinszner, H. and Ron, D.
TITLE Identification of novel stress-induced genes downstream of chop
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE
98315054
2 (bases 1 to 1412)
REFERENCE
AUTHORS Sok, J., Wang, X.Z., Batcharova, N., Kuroda, M., Harding, H. and Ron, D.
TITLE CHOP-dependent stress-inducible expression of a novel form of
carbonic anhydrase VI
JOURNAL Mol. Cell Biol. 19 (1), 495-504 (1999)

MEDLINE 9907987
 LOCUS 3 (bases 1 to 1412)
 REFERENCE Sok,J., Wang,X.Z. and Ron,D.
 AUTHORS Direct Submission
 TITLE Submitted (22-JUL-1998) Skidball Institute, New York University
 JOURNAL Medical Center, 540 First Ave., New York, NY 10016, USA

FEATURES
 source location/Qualifiers
 1..1412
 /organism="Mus musculus"
 /strain="CD1"
 /db_xref="taxon:10090"
 /tissue_type="salivary gland"
 54..125
 54..1007
 /note="alternatively spliced"
 /codon_start=1
 /product="secreted carbonic anhydrase isozyme VI precursor"
 /protein_id="A012540.1"
 /db_xref="GI:3421371"

sig_peptide
 CDS

mat_peptide
 126..1004
 /product="secreted carbonic anhydrase isozyme VI"

BASE COUNT 414 a 343 c 345 g 310 t
 ORIGIN

Query Match 96.2%; Score 511.6; DB 10; Length 1412;
 Best Local Similarity 99.2%; Pred. No. 1.9e-140;
 Matches 514; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

15 AACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAAGTCACTTGTCC 74
 711 AACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAAGTCACTTGTCC 770
 75 AAGGCCCGGTGGACCATGAAACCTGTATGATCAACAACACACATTTCTCA 134
 771 AAGGCCCGGTGGACCATGAAACCTGTATGATCAACAACACACATTTCTCA 830
 135 AATGTTACCGCAGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 194
 831 AATGTTACCGCAGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 890
 195 GTCCAGATATGATCTTCTGTACCACTTAACTAAACATGCAAGAGATTTCTA 254
 891 GTCCAGATATGATCTTCTGTACCACTTAACTAAACATGCAAGAGATTTCTA 950
 255 CAACCTAAGAAACAGAAACAAAGAAAGACCGGACTTTTGGAGCCGAATTCAC 314
 951 CAACCTAAGAAACAGAAACAAAGAAAGACCGGACTTTTGGAGCCGAATTCAC 1010
 315 ACTGCCCCCAGGAGTCACTTGTCTGAAGAAATCTGGAATGTGGTCTCTTG 374
 1011 ACTGCCCCCAGGAGTCACTTGTCTGAAGAAATCTGGAATGTGGTCTCTTG 1070
 375 CTGCTTACAGGAGTCTGTATTAACCCAGAGGAGTCTCTCCGAGAGAAAGTGA 434
 1071 CTGCTTACAGGAGTCTGTATTAACCCAGAGGAGTCTCTCCGAGAGAAAGTGA 1130
 435 GCTAGTTAGCCAGAAAGAGATGTGGTGAAGAGAAAGACTTTCAGCAGAGCTG 494
 1131 GCTAGTTAGCCAGAAAGAGATGTGGTGAAGAGAAAGACTTTCAGCAGAGCTG 1190
 495 CAAGAAATCAAGCTTTCATTAATTCCTGTCATCAT 532
 1191 CAAGAAATCAAGCTTTCATTAATTCCTGTCATCAT 1228

RESULT 2

AF079834 1475 bp mRNA linear ROD 04-FEB-1999
 DEFINITION Mus musculus stress-inducible intra-cellular carbonic anhydrase
 isoform VI mRNA, complete cds.
 AF079834 GI:3421368

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

3 (bases 1 to 1475)
 Sok,J., Wang,X.Z., Batchvarova,N., Kuroda,M., Harding,H. and Ron,D.
 Direct Submission
 Submitted (22-JUL-1998) Skidball Institute, New York University
 Medical Center, 540 First Ave., New York, NY 10016, USA

FEATURES
 source location/Qualifiers
 1..1475
 /organism="Mus musculus"
 /strain="NIH Swiss"
 /db_xref="taxon:10090"
 /cell_line="NIH-3T3"
 285..1070
 /note="alternatively spliced"
 /codon_start=1
 /product="stress-inducible intra-cellular carbonic anhydrase isozyme VI"

CDS

15 AACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAAGTCACTTGTCC 74
 774 AACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAAGTCACTTGTCC 833
 75 AAGGCCCGGTGGACCATGAAACCTGTATGATCAACAACACACATTTCTCA 134
 834 AAGGCCCGGTGGACCATGAAACCTGTATGATCAACAACACACATTTCTCA 893
 135 AATGTTACCGCAGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 194
 894 AATGTTACCGCAGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 953
 195 GTCCAGATATGATCTTCTGTACCACTTAACTAAACATGCAAGAGATTTCTA 254
 954 GTCCAGATATGATCTTCTGTACCACTTAACTAAACATGCAAGAGATTTCTA 1013
 255 CAACCTAAGAAACAGAAACAAAGAAAGACCGGACTTTTGGAGCCGAATTCAC 314
 1014 CAACCTAAGAAACAGAAACAAAGAAAGACCGGACTTTTGGAGCCGAATTCAC 1073

BASE COUNT 423 a 370 c 357 g 325 t
 ORIGIN

Query Match 96.2%; Score 511.6; DB 10; Length 1475;
 Best Local Similarity 99.2%; Pred. No. 1.9e-140;
 Matches 514; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

15 AACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAAGTCACTTGTCC 74
 774 AACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAAGTCACTTGTCC 833
 75 AAGGCCCGGTGGACCATGAAACCTGTATGATCAACAACACACATTTCTCA 134
 834 AAGGCCCGGTGGACCATGAAACCTGTATGATCAACAACACACATTTCTCA 893
 135 AATGTTACCGCAGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 194
 894 AATGTTACCGCAGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 953
 195 GTCCAGATATGATCTTCTGTACCACTTAACTAAACATGCAAGAGATTTCTA 254
 954 GTCCAGATATGATCTTCTGTACCACTTAACTAAACATGCAAGAGATTTCTA 1013
 255 CAACCTAAGAAACAGAAACAAAGAAAGACCGGACTTTTGGAGCCGAATTCAC 314
 1014 CAACCTAAGAAACAGAAACAAAGAAAGACCGGACTTTTGGAGCCGAATTCAC 1073

Matches	327	Conservative	0	Mismatches	3	Indels	0	Gaps	0
OY	203	TATGTACTCTTCGTGACACACTTTACCTAAAAACATGCAGAAAGGAGTTCTACAACCTTAA	262						
Db	9470	TCTATACACTTTCGTACCCACTTTACTTACCTAAAAACATGCAGAAAGGAGTTCTACAACCTTAA	9411						
OY	263	GAACAAGAAAGAAACAAAGAAAGAACCCGCACCTTTTGAGACCCGGAATATACACACTGSGCC	322						
Db	9410	GAACAAGAAAGAAACAAAGAAAGAACCCGCACCTTTTGAGACCCGGAATATACACACTGSGCC	9351						
OY	323	CAGGGCTACCCCTGGCCCTTGTCTAAGGAATCTTGGAATGTGGGTCTCTCTGCTGCCCTTC	382						
Db	9350	CAGGGCTACCCCTGGCCCTTGTCTAAGGAATCTTGGAATGTGGGTCTCTCTGCTGCCCTTC	9291						
OY	383	AGCGAGTCTCTGATTAACCCAGAGGAGTCTCTCCGACAGAAAGATGAGACTTACGTT	442						
Db	9290	AGCGAGTCTCTGATTAACCCAGAGGAGTCTCTCCGACAGAAAGATGAGACTTACGTT	9231						
OY	443	AGCCAGAGAGAAAGGAGTGTGGTGGAAGAGGAAAGACTTCGACGAGCTGCAAAAGAA	502						
Db	9230	AGCCAGAGAGAAAGGAGTGTGGTGGAAGAGGAAAGACTTCGACGAGCTGCAAAAGAA	9171						
OY	503	TCAAGCTTTTATTAATTCCTGTCATCAT	532						
Db	9170	TCAAGCTTTTATTAATTCCTGTCATCAT	9141						
RESULT 4									
AL606971/c									
LOCUS	AL606971	226957 bp	DNA	linear	HTG 15-JAN-2002				
DEFINITION	Mus musculus chromosome 4 clone RP23-123120, *** SEQUENCING IN								
ACCESSION	AL606971								
VERSION	AL606971.6	GI:18181715							
KEYWORDS	HTG; HTGS.PHASE1; HTGS.ACTIVEFIN; HTGS_DRAFT; HTGS_FUULLTOP.								
SOURCE	house mouse.								
ORGANISM	Mus musculus								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
REFERENCE	1 (sites)								
AUTHORS	Hunt,A.								
TITLE	Direct Submission								
JOURNAL	Submitted (14-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,								
	Cambridgeshire, CB10 1SA, UK. E-mail enquiries:								
	humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk								
	on Jan 16, 2002 this sequence version replaced gi:15865162.								
COMMENT	----- Genome Center								
	Center: Wellcome Trust Sanger Institute								
	Center code: SC								
	Web site: http://www.sanger.ac.uk								
	Contact: humquery@sanger.ac.uk								
	----- Project Information								
	Center project name: BM123120								
	----- Summary Statistics								
	Assembly program: XGAP4; version 4.5								
	Sequencing vector: plasmid; I08752; 100% of reads								
	Chemistry: Dye-terminator Big Dye; 100% of reads								
	Consensus quality: 224787 bases at least Q40								
	Consensus quality: 225129 bases at least Q30								
	Consensus quality: 225417 bases at least Q20								
	Insert size: 225857; sum-of-contigs								
	Insert size: 203349; 5.9% error; agarose-tp								
	Quality coverage: 8.34x in Q20 bases; sum-of-contigs quality								
	coverage: 9.61x in Q20 bases; agarose-tp								

	* NOTE: This is a 'working draft' sequence.								
	* This record will be updated with the finished sequence								
	* as soon as it is available and the accession number will								
	* be preserved.								
FEATURES									
Source	Location/Qualifiers								
	1..226957								
	/organism="Mus musculus"								
	/db_xref="taxon:10090"								

		/chromosome="4"	
		/clone="RP23-123I20"	
		/clone.lib="RPCR-23"	
		I. .39687	
		/note="assembly_fragment:05052	
		fragment_chain:1	
		clone_end:SP6	
		vectorside:left"	
		39988..58240	
		/note="assembly_fragment:00374	
		fragment_chain:1"	
		58341..65122	
		/note="assembly_fragment:03325	
		fragment_chain:1"	
		65223..68735	
		/note="assembly_fragment:01125	
		fragment_chain:1"	
		68836..81936	
		/note="assembly_fragment:01677	
		fragment_chain:2"	
		82037..101775	
		/note="assembly_fragment:01934	
		fragment_chain:2"	
		101876..120162	
		/note="assembly_fragment:04254	
		fragment_chain:2"	
		120263..160511	
		/note="assembly_fragment:00932	
		fragment_chain:2"	
		160612..167303	
		/note="assembly_fragment:03207	
		fragment_chain:2"	
		167404..206353	
		/note="assembly_fragment:04640	
		fragment_chain:2"	
		206454..211015	
		/note="assembly_fragment:03065	
		fragment_chain:3"	
		211116..226957	
		/note="assembly_fragment:03747	
		fragment_chain:3"	
BASE COUNT	56050 a 55247 c 55829 g 58715 t 1116 others		
ORIGIN			
Query Match	61.1%; Score 325.2; DB 2; Length 226957;		
Best Local Similarity	99.1%; Pred. No. 5.8e-85;		
Matches 327; Conservative	0; Mismatches 3; Indels 0; Gaps 0;		
OY 203	TATGTACTCTTCGTAACCACTTACTACAAAAACAATGCAGAGGAGATTCTACAACTNA	262	
DB 118729	TCTAGACTCTTGATGCCACTTTACTATAAAAACATGCGAAGAGATTCTTCAACCTNA	118670	
OY 263	GAAACAGAAGAAACAAAGAACACGGCACATTTTGAGACCAGGAATGACACACTGGCCC	322	
DB 118669	GAAACAGAAGAAACAAAGAACACGGCACATTTTGAGACCAGGAATGACACACTGGCCC	118610	
OY 323	CAGGGCTCACCCTCCCTCTGTCTAAGGAATCCTGGAATGTGGGTCTCTTGCTGCTTC	382	
DB 118609	CAGGGCTCACCCTCCCTGTCTAAGGAATCCTGGAATGTGGGTCTCTTGCTGCTTC	118550	
OY 383	AGCGAGTCTGTATTACCCAGAGGAGATTCTCTCCGAGAGAGAAAGTGGAGCTACGTT	442	
DB 118549	AGCGAGTCTGTATTACCCAGAGGAGATTCTCTCCGAGAGAGAAAGTGGAGCTACGTT	118490	
OY 443	AGCCAGAAAGAGAGGATGTGGGTGAGAGGGAAGAAAGACTTGACGGACGTGCAGAAAGAA	502	
DB 118489	AGCCAGAAAGAGAGGATGTGGGTGAGAGGGAAGAAAGACTTGACGGACGTGCAGAAAGAA	118430	
OY 503	TCAAAGCTTTCAATATCCCTGCTCAATCAT	532	
DB 118429	TCAAAGCTTTCAATATCCCTGCTCAATCAT	118400	

LOCUS	1316 bp	MRNA	linear	PRI 31-OCT-1994
DEFINITION	Human carbonic anhydrase isozyme VI (CA6)	MRNA, complete cds.		
ACCESSION	M57892.1	J05305		
VERSION	M57892.1	G1:119731		
KEYWORDS	carbonate dehydratase; carbonic anhydrase isozyme VI; secreted protein.			
SOURCE	Human salivary gland, cDNA to mRNA.			
ORGANISM	Homo sapiens			
REFERENCE	Maniatis, T. et al. (1982) Molecular cloning, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.			
AUTHORS	Maniatis, T. et al. (1982)			
TITLE	Human secreted carbonic anhydrase: cDNA cloning, nucleotide sequence, and hybridization histochemistry			
JOURNAL	Biochemistry			
MEDLINE	91105141			
FEATURES	Location/Qualifiers			
source	1..1316			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="1p36"			
	/tissue_type="salivary gland"			
	<1..1316			
mrna	/gene="CA6"			
	/note="G00-125-350; putative"			
	1..1316			
gene	/gene="CA6"			
	7..933			
	/EC_number="4.2.1.1"			
	/note="putative"			
	/codon_start=1			
	/product="carbonic anhydrase isozyme VI"			
	/protein_id="AA51892.1"			
	/db_xref="GI:179732"			
	/db_xref="GDB:G00-125-350"			
	/translation="MRAVLVLLSLFLGGQAQHYSDMTYSEGLDEAHMPPHNTYRACGG ORSDINMORTKRYNPSLGLNNTGTEIGAGFPMVNGHTVQIGLPSTRMRTVADG IVYIAQOMHFHWGASSELISGSEHTVQIRHVEIHIVYNSKRYKIDQAPDGLA VLAEFVNKNPENTYYSNFIASHLANKYRGQRTTLGLVDMLPRLDLYHYTHSS LTPRECTENVMHFLADPVKLSRQWKLKLESLDHNKTIINDYRTPQYLKNHVES NFPNQEYTLGSEFPQYFLKRIEILIDYLRALN"			
	7..57			
sig_peptide	/gene="CA6"			
	/note="G00-125-350; putative"			
	58..930			
mat_peptide	/gene="CA6"			
	/EC_number="4.2.1.1"			
	/note="G00-125-350; putative"			
	/product="carbonic anhydrase isozyme VI"			
BASE COUNT	362 a 343 c 313 g 298 t			
ORIGIN	Chromosome 1.			
Query Match	19.1%; Score 101.8; DB 9; Length 1316;			
Best Local Similarity	67.8%; Pred. No. 4.4e-19;			
Matches 158; Conservative	0; Mismatches 72; Indels 3; Gaps 1;			
15	ACACCCCCCTGCACAGAGATGTCCAGTGTGTGCTTAGACACAAAGTCACCTTGTCC 74			
667	ACGCTCTCCTGCACAGAGATGTCCAGTGTGTGCTTAGACACAAAGTCACCTTGTCC 726			
75	AAGGCCACAGTGTGATCCATGAAAACCTGTGTAAGATGACCAACACACCATTCGA 134			
727	AGGACACAGTGTGATCCATGAAAACCTGTGTAAGATGACCAACACACCATTCGA 786			
135	AATGGTTACCGCAGACACACCCCAACACACAGAGGTGTGGAGCCCAATTTCTGTA - 192			
787	AACGATTAACGAGAGACCCGCTGAAACACAGAGTGTGGATTCACACTTCCCGAAT 846			
193	-ACGTCCAGGATATGTAATCTTCTTGTACCAACCTTTACTCTTAAAAACATGCAGAA 244			

```
FEATURES
source      location/qualifiers
1.107714   /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/amp="p36.21-36.33"
```

FEATURES
SOURCE

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30): an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TRMBL; Wp., WORMPEP. Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/Ref/chr1>

RP3-477M7 is from the library RP01-3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP3-477M7.

repeat_region	293. .435	/note="AluJo repeat: matches 1. .140 of consensus"
repeat_region	436. .747	/note="AluSg repeat: matches 1. .310 of consensus"
repeat_region	748. .894	/note="AluJo repeat: matches 140. .297 of consensus"
repeat_region	994. .1052	/note="L1ME3 repeat: matches 5987. .6047 of consensus"
repeat_region	1089. .1264	/note="RHIC repeat: matches 1. .133 of consensus"
repeat_region	1265. .1577	/note="AluSx repeat: matches 2. .311 of consensus"
repeat_region	1578. .1800	/note="RHIC repeat: matches 133. .371 of consensus"
repeat_region	1941. .2213	/note="AluSx repeat: matches 21. .293 of consensus"
repeat_region	2353. .2454	/note="MIR repeat: matches 80. .172 of consensus"
misc_feature	/note="complement(2444. .2843)	
misc_feature	/note="match: GSS: Em:AQ076041"	
misc_feature	/note="match(2455. .2645)	
misc_feature	/note="match: GSS: Em:AQ075536"	
misc_feature	complement(2469. .2684)	
misc_feature	/note="match: GSS: Em:AQ065754"	
misc_feature	complement(2586. .2904)	
misc_feature	/note="match: GSS: Em:AQ005352"	
repeat_region	2794. .2958	/note="MER5B repeat: matches 1. .178 of consensus"
repeat_region	2979. .3115	/note="F1AM_A repeat: matches 1. .125 of consensus"
repeat_region	3141. .3430	/note="AluSx repeat: matches 1. .290 of consensus"
repeat_region	4013. .4159	/note="HERVL repeat: matches 5615. .5764 of consensus"
repeat_region	4162. .4370	/note="MIR2E repeat: matches 13. .188 of consensus"
repeat_region	4321. .4614	/note="AluSx repeat: matches 1. .296 of consensus"
repeat_region	4615. .4706	/note="MIR2E repeat: matches 188. .275 of consensus"
repeat_region	4790. .4944	/note="MIR2E repeat: matches 130. .286 of consensus"

Query Match 12.6%; Score 67; DB 9; Length 432;
Best Local Similarity 73.9%; Pred. No. 7,7e-09;
Matches 85; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

OY 79 CCCAGTGTGACCATAGAAACTCTGTATGCATCAACAACACCATTCATAATG 138
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 279 CCCAGGTTGGAAAGCTGAGAAATCTTACTGATCACCCAGACAGACCATCCACG 338
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
OY 139 GTTACCCGACACACACACACACACAGCGGTGTGTAAGCAATTCCTGAA 193
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 339 ATTACCGCAGACCCAGCCCTGTAAACACAGAGTGTGTAATCCACTTCCGAA 393
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT 10
AC091984/1 188036 bp DNA linear HTG 09-JUN-2001
LOCUS AC091984/1
DEFINITION Homo sapiens chromosome 5 clone RP11-569B13, WORKING DRAFT
SEQUENCE AC091984
AC091984.1 GI:14333920
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 188036)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188036)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web Site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 624898
Center clone name: RPCI-11_569B13
-----
Summary Statistics
Consensus quality: 149093 bases at least Q40
Consensus quality: 171742 bases at least Q30
Consensus quality: 175083 bases at least Q20
Estimated insert size: 174490; agarose-fp estimation
Estimated insert size: 185136; sum-of-contigs estimation
Quality coverage: 6.47 in Q20 bases; agarose-fp estimation
Quality coverage: 6.05 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1488: contig of 1488 bp in length
* 1489 1588: gap of unknown length
* 1589 2840: contig of 1252 bp in length
* 2841 2940: gap of unknown length
* 2941 4281: contig of 1341 bp in length
* 4282 4382: gap of unknown length
* 4382 5562: contig of 1181 bp in length
* 5562 5663: gap of unknown length
* 5663 6777: contig of 1115 bp in length
* 6777 6878: gap of unknown length
* 6878 7936: contig of 1059 bp in length
* 7937 8036: gap of unknown length
* 8037 9106: contig of 1070 bp in length
* 9107 10483: contig of 1277 bp in length
* 10484 10583: gap of unknown length
* 10584 12565: contig of 1982 bp in length

```

```

* 12566 12665: gap of unknown length
* 12666 15417: contig of 2752 bp in length
* 15418 15517: gap of unknown length
* 15518 20425: contig of 4908 bp in length
* 20426 20525: gap of unknown length
* 20526 25068: contig of 4543 bp in length
* 25069 25168: gap of unknown length
* 25169 30263: contig of 5095 bp in length
* 30264 30363: gap of unknown length
* 30364 34974: contig of 4611 bp in length
* 34975 35074: gap of unknown length
* 35075 43270: contig of 8196 bp in length
* 43271 43371: gap of unknown length
* 43371 49361: contig of 5991 bp in length
* 49362 49461: gap of unknown length
* 49462 56288: contig of 6827 bp in length
* 56289 56389: gap of unknown length
* 56389 63723: contig of 7335 bp in length
* 63724 63823: gap of unknown length
* 63824 69752: contig of 5929 bp in length
* 69753 69852: gap of unknown length
* 69853 78189: contig of 8337 bp in length
* 78190 78289: gap of unknown length
* 78290 85925: contig of 7636 bp in length
* 85926 86025: gap of unknown length
* 86026 92953: contig of 6528 bp in length
* 92954 93053: gap of unknown length
* 93054 101358: contig of 8305 bp in length
* 101359 101458: gap of unknown length
* 101459 108969: contig of 7511 bp in length
* 108970 109069: gap of unknown length
* 109070 117359: contig of 8280 bp in length
* 117360 117459: gap of unknown length
* 117460 126188: contig of 8729 bp in length
* 126189 126288: gap of unknown length
* 126289 138034: contig of 11746 bp in length
* 138035 138134: gap of unknown length
* 138135 150053: contig of 11919 bp in length
* 150054 150153: gap of unknown length
* 150154 169677: contig of 19424 bp in length
* 169678 169678: gap of unknown length
* 169679 188036: contig of 18359 bp in length.
* Location/Qualifiers
  source          1..188036
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="5"
                  /clone="RP11-569B13"
                  /clone_1lb="RPCI human BAC library 11"
                  /clone_1lb="RPCI human BAC library 11"
BASE COUNT      57018 a 36648 c 33629 g 57817 t 2924 others
ORIGIN
Query Match      8.0%; Score 42.4; DB 2; Length 188036;
Best Local Similarity 66.3%; Pred. No. 0.32;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
OY 219 CACCTTACTAATAAACAAGCAGAGATCTCAACCTAAGAAAGAGAAACA 278
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 61874 CACCTGTATTGAATAAAGAAAGACTGACATCTTCAATTGAGGCAAGAAACA 61815
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
OY 279 AAGAAGACCGGACACTTTGGAGCCGGAATG 310
    ||  ||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 61814 AATGGAACTAGCATTTTGAACAGGAACTG 61783
    ||  ||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT 11
AF165124 195909 bp DNA linear PRI 07-JUN-2001
LOCUS AF165124
DEFINITION Homo sapiens chromosome 5q31.1-q33.1 clone BAC djn082c10 containing
GABRG2 gene, complete sequence.
ACCESSION AF165124
VERSION AF165124.1 GI:5738137
KEYWORDS HTG.

```


ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 209836)
JOURNAL	DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE	Direct Submission
AUTHORS	Unpublished
TITLE	2 (bases 1 to 209836)
JOURNAL	DOE Joint Genome Institute.
REFERENCE	Direct Submission
AUTHORS	Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint
TITLE	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
JOURNAL	3 (bases 1 to 209836)
REFERENCE	DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS	Direct Submission
TITLE	Submitted (30-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
JOURNAL	Drive, Walnut Creek, CA 94598, USA
REFERENCE	4 (bases 1 to 209836)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
COMMENT	Drive, Walnut Creek, CA 94598, USA On Sep 21, 2001 this sequence version replaced g1:15375175. Draft Sequence produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing completed at Stanford Human Genome Center www.shgc.stanford.edu Quality: Phrap Quality >=40 99.8% of Sequence; Estimated Total Number of Errors is 0.3. STS Content: SHGC-79397 G52187.
FEATURES	
SOURCE	Location/Qualifiers 1..209836 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="5" /clone="RP11-290C6"
BASE COUNT	66857 a 39092 c 37729 g 66158 t
ORIGIN	
Query Match	8.0%; Score 42.4; DB 9; Length 209836;
Best Local Similarity	66.3%; Pred. No. 0.33;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;	
QY 219 CACCTTACTTAATAAACATCGAGAAAGGATGTTCACCACTTAGAAGAAGAAAACA 278	
Db 65753 CACOTGTATTGAANAATAATGAACACATCTCTTCATATTGAGGGCAAAGAAAAACA 65812	
OY 279 AAGAAGACCAGCATTTTGAGACCCGGAATG 310	
Db 65813 AATGGAGACTAGCATTTTTTGAACGAGAACTG 65844	
RESULT 13	
AE004453	14568 bp DNA linear BCT 30-AUG-2000
LOCUS	Pseudomonas aeruginosa PAO1, section 14 of 529 of the complete
DEFINITION	genome.
ACCESSION	AE004453 AE004091
VERSION	AE004453.1 GI:9945978
KEYWORDS	.
SOURCE	Pseudomonas aeruginosa.
ORGANISM	Pseudomonas aeruginosa Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.
REFERENCE	1 (bases 1 to 14568) Stover,C.K., Pham,X.O., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.J., Goltry,L., Tolentino,E., Westbrook-Medman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Laidig,K., Lim,K., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T. Complete genome sequence of Pseudomonas aeruginosa PAO1, an
TITLE	

JOURNAL Nature 406 (6799), 959-964 (2000)

MEDLINE 20437337

REFERENCE 2 (bases 1 to 14568)

AUTHORS Storer, C.K., Pham, X.-O.T., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Huftagel, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltzy, L., Tolentino, E., Westbrock-wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Lardig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Sailer, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.

TITLE Direct Submission

JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

FEATURES

source 1. 14568

organism="Pseudomonas aeruginosa"

strain="PA01"

/db_xref="taxon:287"

83. .1033

/gene="PA0148"

83. .1033

/gene="PA0148"

/codon_start=1

/transl_table=11

/product="probable adenosine deaminase"

/protein_id="AG03538.1"

/db_xref="GI:9945979"

/translation="MYEWLNLATPKALHLHLEGTLEPELLEFALAERNRIATLFWNDVET LKAVYNNLOEFLDYAGADVLRTEODPYDLTWYVLOKCAONVHVEPEFDQTH TDGRTPEVYLAIRALRNGEKLIGIRGLITLSPRLSEDAOKLDOALPFRFAD IANGLDSEVGHPPSKFQRYFEDPABSEGLTVNAHGEPPPEYIWEALDLLVERDH GVAEFDELRMLRLIDEQIPLYVCPISNTKLCYFDDMSQHTILDMERKVTYVSD PAYFGVYTFENFHALQOSLGMEEQARLQNLSLARLVK"

1231. .1776

/gene="PA0149"

1231. .1776

/gene="PA0149"

/codon_start=1

/transl_table=11

/product="probable sigma-70 factor, ECF subfamily"

/protein_id="AG03539.1"

/db_xref="GI:9945980"

/translation="MPADGEARRADLLKAFLAOROMESLVSRYGCRATADLVQE LFLRFRPPAPVPEIGSLYSRLSARNALIDHLEGSROROLDWLPPEOROGPEVPE QALBAGDQMRVREALRGDLPERRRITFLNRHGRYAEIAQAMQLSGSAVEKHMARA LDACKASLAGPAGEQARSQP"

1773. .2768

/gene="PA0150"

1773. .2768

/gene="PA0150"

/codon_start=1

/transl_table=11

/product="probable transmembrane sensor"

/protein_id="AG03540.1"

/db_xref="GI:9945981"

/translation="WTRPSDPAAYEPVLEQTAQAWLVRLGTPDAATRRAFEDWLAQ RPAPAPAYARAEALMRLTEIPASRLAIEADALEAYLPARRASNPRRMAYPLASVAC LLAMLMGGMQWQPGDWLQDMRDADVYVAPGOVAEMKLSLIDAGSAVRVLDGGO RRVLELRGAALRVKRDALPVPVAGGSRARVLTETFEYRRGAGATPVYVLOGRVAYS AGTGAGAAEEVLTGAGQOVGFADGKLRPVADAAALAMRQGMVSTFRRLAEVYDEL ARVYPRITLLDDALGRQPVGSFRSDDPPEALAKSLQAVLGRQQLLEGRLVVIR"

2917. .5304

/gene="PA0151"

2917. .5304

/gene="PA0151"

/codon_start=1

/transl_table=11

/product="probable TonB-dependent receptor"

/protein_id="AG03541.1"

/db_xref="GI:9945982"

/translation="WTPPRRALRTIPALFISGLCIAATVQAAEAVVRFSQAAAPLA QALNFRASQOTLITVETLPPQATAPALHGSIVAEALQRLDLOGSLRRARIDERLT

TLPEIPASGALDATTVSAAAPQCYRPAANASISRGDPAVLEIPQAI5VPAQALQ DORPNLIDALINGISGTOANTIGTQDAMKRGEDNDGDSIMRGMP5VQGRNETA TADHVEIKPAPSLIYGIDPGGVVNVVSKKPDLOQANALTLRGSAYAHARGSGGOL DSTGPDLSLAIKMLIDHDEDEYWRDPOKSRPTETIAPSLAWGEEDTVNLYEHEEF TTPEDRQTAIDPPTNRPLDIPRSRLDEPNITEGSELTRLDIPEOLDIPKHAHGY GYSRETYDDQARVYVMAVDGTLTRRDGTGHAVSSDFTTGLSETLOLAGMHDQ FGMHDEKRIERADILDRDSKTSREPYLDVYGGVSTPVSIVAVDSDQTKLSDSLIFE ODSIHLEDERILVIGARVQIYDQLAGRGREFNANDINQKAVPRGGLVYRUSDEVSL YGSMRSFKNSNTIAPISGETTIDSAIPEEATSMELGAKLDVPGRLSGTLALFIRK KNVLYNLDGANSNSYAPAGRSRGLIEDITGOLSEYRSLISYSYAMLAAYEDEDTL EGRKLVNARKTSLASVYELGOLFPGDDRIRLGGGARYGKRGPDANSFEDLPSTYVA DAFASYDTRFEGGKVFQLVNKNLFDRTYTPSSANRLYVAMGPRFOVSTYEE"

5681. .6508

/gene="pcaQ"

/note="PA0152"

5681. .6508

/gene="pcaQ"

/codon_start=1

/transl_table=11

/product="transcriptional regulator pcaQ"

/protein_id="AG03542.1"

/db_xref="GI:9945983"

/translation="MAVSOPAI5STLKELELIGARLFERSKAGAEITENGVTFLRYA GPVQALRDGYNITRGHDAQVETRYRGVLSVESIYVPEYIRLRHRRHGLVYSVANG PSAYLQALRVGELDLYVGRWTDSPDIOGLTFEHLSESMALVYRGHPILAPLGLDL QRLAEFLVPLPLAGTTIRHADSLFVQWGLALGOQRLLETISPLTSRYVRSATIVYA PRDAVRIDVNGCELAELDLGIREPGSVGLCHNAALPLPLGAQWFSVILREVAELH"

6643. .7362

/gene="pcaH"

/note="PA0153"

6643. .7362

/gene="pcaH"

/codon_start=1

/transl_table=11

/product="3,4-dioxygenase, beta subunit"

/protein_id="AG03543.1"

/db_xref="GI:9945984"

/translation="WHDASRRVYTRDRNMHPKALTPDYKTSVRSPOALVSPSL SETSGPDSHLRLGHDNDLLNFDHGLPGGERTIMRGYFDQYGRPVHTLVEMQD ANAGRYRHKNDRYLAPLDPNFGVGLTDESEHYFRITKEGYPWRGNPMWRPA HIIHVSQSPATIAARLYATQMYFEGDPLIPKCPILITLADPEAAOSLIGRLDSMANPMD CLAYRFDVILRGORTPEPENA"

7373. .7978

/gene="pcaG"

/note="PA0154"

7373. .7978

/gene="pcaG"

/codon_start=1

/transl_table=11

/product="3,4-dioxygenase, alpha subunit"

/protein_id="AG03544.1"

/db_xref="GI:9945985"

/translation="MSHELRLRTPSQTAGPVYVHIGLALAANGNPTRLDELTNRMRARG ARGQHILLGHVYDNGHIVRDSFLEFWQDSSEGYQORYPERPFNSFGRTATFDG GEMTLHTVAGVYRNAQGVPMAPHINVSILFGRINAIHQITRLYEDDETEANAACPVNL ILEQAPRRDTLLAQRCEVDQALAYREFDIRIOGDNTEVFDE"

8185. .9024

/gene="pcaR"

/note="PA0155"

8185. .9024

/gene="pcaR"

/codon_start=1

/transl_table=11

/product="transcriptional regulator pcaR"

/protein_id="AG03545.1"

/db_xref="GI:9945986"

/translation="MSELEPHPTLAPPTVLSPARIRAFTEGDPNMTSLARGLAVIH AFOERKRLHTIAOISHREIIPRAAVRCHLTMOGVTATDORISLIPKVLTGSHAY ISSPTIAPAOPIIDRLSROLHEACSMATLEBDVLYTARSATPQRLISVDLVNWSRL PAYCTSMGIIILAAIDDDALHAIVFGVEQAKTSRTLYTPETILCLIVEIRROGCIY DDELEGLRSIAVPVADSAGHYLAALNVTGTHAGVSRABLESRLPLLELSRELSAR LFP"

9177. .10328

```

CDS
/gene="PA0156"
/177..10328
/gene="PA0156"
/codon_start=1
/transl_table=11
/product="probable RND efflux membrane fusion protein precursor"
/protein_id="AAG03546.1"
/db_xref="GI:9945987"
/translation="MSDARGAFHSKGRMSRMLPA1LCAQHLVCGGAAPPAEHHVRL
AQTVAKEFASATITGDIQARVADQSFVGGKIVERTLVGDVHVAAGVLAFLDPO
DORSNENQAQAAQAAQSLADLNTYORAKLIPKVTSGSEVDALASVRSQSSLI
KAAQALANARDLSTETELASDAGVITARAQAEVGVQVAVPTTLARQDEBDAVN
VYSELSHDVDCORITVSLKPEVTASGVREITPTVDESGLTKVAGVDSVPAEM
SLGSVYMASVAPAEHSHVTLPSASLSTKGEQPAWLLDQOKARLQVRAVASEKY
VIDGLEGAGQVTVTGSGDLHFGVVEVAQPPQSTASDAVGGGP"
10325..11395
/gene="PA0157"
/10325..11395
/gene="PA0157"
/codon_start=1
/transl_table=11
/product="probable RND efflux membrane fusion protein precursor"
/protein_id="AAG03547.1"
/db_xref="GI:9945988"
/translation="MKPFSLAGLGFALLSGCDEDEPPAPPPVLTMYTKIKNDL
GRFASIQARESVYGFRTNGRIASRLFDGDFGKALLATIDPDQOQILASQSD
LASAEQIDIAQANAROEELFARSVAOARIDARPLRTSQASFDQAKAAVQQRD
OLSTRLVTDGDIQARVADQSFVGGKIVERTLVGDVHVAAGVLAFLDPO
ARFLVSAQDLQARVADQSFVGGKIVERTLVGDVHVAAGVLAFLDPO
AGRSVTPASVTLERQKTOVWVVDKSGVALREVQVLSRDERQVVGQGLADGDR
VVRAGVNSLAKGQKIKLDEDA"
11392..14439
/gene="PA0158"
/11392..14439
/gene="PA0158"
/11392..14439

Query Match
Best Local Similarity 55.6%; Pred. No. 0.4;
Matches 80; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 63 GTCACTTTGCCAAGCCAGGTGACATGACAAACCTGTTATGATGATCAACACAC 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12484 GTGGCTCTCGATGCCCGCTGCTGCGCATGCTCTTCTGCTTACATGATACACCGAC 12543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 AACACCATTCAAATGTTTACCGACACACACCCAAACACACAGGCTGTGAAGCC 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12544 ATCACCATGCGACGGGGTTTCCCTCGCGCGCTGATCATCGCCCTCGGCTGTGTGGAC 12603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 183 AATTTCTGAACGTCCAGGATATG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12604 GATGCATGATCACCCTGGAGATG 12627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
LOCUS AF128411S6 1077 bp DNA linear PRI 02-JAN-2000
DEFINITION Homo sapiens carbonic anhydrase VI gene, exon 6.
ACCESSION AF128416
VERSION AF128416.1 GI:6652972
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1077)
AUTHORS Grubb, D.J.
JOURNAL Direct Submission
Submitted (16-FEB-1999) Howard Florey Institute, University of
Melbourne, Royal Parade, Melbourne, Vic 3133, Australia
FEATURES
source
1..1077
Location/Qualifiers

```

```

BASE COUNT 239 a 309 c 228 g 301 t
ORIGIN
exon
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.2"
/number=6

Query Match
Best Local Similarity 72.0%; Pred. No. 0.33;
Matches 54; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 15 ACACCCCGCGACAGATGCGCTGCTTGTGCTTACGACAAAGTCATTTGCC 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 ACAGCTCCCGACCTAGACACGCTGCTGCTGCGAGATTTGCAAGCTTCC 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 75 AAGCCCGAGCTGTG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 666 AGGACACAGGTATG 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
LOCUS AC098147
DEFINITION Rattus norvegicus clone CH230-173010, *** SEQUENCING IN PROGRESS -
ACCESSION AC098147
VERSION AC098147.2 GI:17975713
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 82576)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaralunga, H.C., Aze, J.R., Banks, T., Barbara, J.,
Belton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowle, S., Brieval, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C.,
Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-carroll, L., Dederich, D.A., Delanaye, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
Hollins, B., Homsl, F., Howard, S., Huber, J., Hui, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewise, L., Li, J., Li, Z., Lichteberg, O., Lieu, C., Liu, J., Liu, W.,
Lounseged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Kapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabadi, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S.,
Ogih, M., Okunolu, G., Oragunye, N., Oyedro, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshari, N.,
Sisson, I., Sodergren, E., Sonakke, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Teiford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Wlezyk, R., Wooden, S.,
Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Morley, K., Wu, C., Wu, Y., Zhou, J., Zorilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 82576)

Worley, K. C.

Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 21, 2001 this sequence version replaced gi:16327853.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GGMU

Center clone name: CH230-173010

Summary Statistics

Assembly program: Phrap; version 0.990329first call to

findPhrapList

Consensus quality: 70512 bases at least Q40

Consensus quality: 74282 bases at least Q20

Consensus quality: 76952 bases at least Q20

Estimated insert size: 40288; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 0.5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 54 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1970: contig of 1970 bp in length
1 1971 2070: gap of unknown length
* 2071 3757: contig of 1687 bp in length
* 3758 3857: gap of unknown length
* 3858 5420: contig of 1563 bp in length
* 5421 5520: gap of unknown length
* 5521 6870: contig of 1350 bp in length
* 6871 6970: gap of unknown length
* 6971 9455: contig of 2485 bp in length
* 9456 9555: gap of unknown length
* 9556 11267: contig of 1712 bp in length
* 11268 11367: gap of unknown length
* 11368 13657: contig of 2290 bp in length
* 13658 13757: gap of unknown length
* 13758 15356: contig of 1599 bp in length
* 15357 15456: gap of unknown length
* 15457 16787: contig of 1331 bp in length
* 16788 16887: gap of unknown length
* 16889 18915: contig of 2028 bp in length
* 18916 19015: gap of unknown length
* 19016 20781: contig of 1766 bp in length
* 20782 20881: gap of unknown length
* 20882 22102: contig of 1221 bp in length
* 22103 22202: gap of unknown length
* 22203 23935: contig of 1733 bp in length
* 23936 24035: gap of unknown length
* 24036 26278: contig of 2243 bp in length
* 26279 26378: gap of unknown length
* 26379 27526: contig of 1148 bp in length
* 27527 27626: gap of unknown length
* 27627 29888: contig of 2262 bp in length
* 29889 31664: gap of unknown length
* 31665 31764: contig of 1676 bp in length
* 31765 31764: gap of unknown length

31765 33071: contig of 1307 bp in length
* 33072 33171: gap of unknown length
* 33172 34578: contig of 1407 bp in length
* 34579 34678: gap of unknown length
* 34679 36329: contig of 1651 bp in length
* 36330 36429: gap of unknown length
* 36430 37667: contig of 1237 bp in length
* 37668 37766: gap of unknown length
* 37767 38839: contig of 1073 bp in length
* 38840 38939: gap of unknown length
* 38940 40694: contig of 1755 bp in length
* 40695 40794: gap of unknown length
* 40795 42235: contig of 1441 bp in length
* 42236 42335: gap of unknown length
* 42336 42947: contig of 1512 bp in length
* 42948 43848: gap of unknown length
* 43849 45203: contig of 1255 bp in length
* 45204 45302: gap of unknown length
* 45303 46485: contig of 1183 bp in length
* 46486 46585: gap of unknown length
* 46586 47935: contig of 1350 bp in length
* 47936 48035: gap of unknown length
* 48036 49282: contig of 1246 bp in length
* 49283 49381: gap of unknown length
* 49382 50525: contig of 1144 bp in length
* 50526 50625: gap of unknown length
* 50626 52144: contig of 1519 bp in length
* 52145 52244: gap of unknown length
* 52245 53714: contig of 1470 bp in length
* 53715 53814: gap of unknown length
* 53815 54647: contig of 1033 bp in length
* 54648 54947: gap of unknown length
* 54948 56229: contig of 1282 bp in length
* 56230 56329: gap of unknown length
* 56330 57389: contig of 1059 bp in length
* 57389 57488: gap of unknown length
* 57489 59042: contig of 1554 bp in length
* 59043 59142: gap of unknown length
* 59143 60524: contig of 1382 bp in length
* 60525 60624: gap of unknown length
* 60626 61854: contig of 1230 bp in length
* 61855 61954: gap of unknown length
* 61955 62998: contig of 1044 bp in length
* 62999 63098: gap of unknown length
* 63099 64589: contig of 1491 bp in length
* 64590 64689: gap of unknown length
* 64690 65894: contig of 1205 bp in length
* 65895 65994: gap of unknown length
* 65995 67173: contig of 1179 bp in length
* 67174 67273: gap of unknown length
* 67274 68458: contig of 1184 bp in length
* 68459 68557: gap of unknown length
* 68558 69646: contig of 1089 bp in length
* 69647 69746: gap of unknown length
* 69747 70862: contig of 1116 bp in length
* 70863 70962: gap of unknown length
* 70963 71978: contig of 1016 bp in length
* 71979 72078: gap of unknown length
* 72079 73487: contig of 1408 bp in length
* 73488 73586: gap of unknown length
* 73587 74589: contig of 1003 bp in length
* 74590 74689: gap of unknown length
* 74690 76223: contig of 1534 bp in length
* 76224 76323: gap of unknown length
* 76324 77585: contig of 1262 bp in length
* 77586 77685: gap of unknown length
* 77686 78798: contig of 1113 bp in length
* 78799 78899: gap of unknown length
* 78900 80072: contig of 1174 bp in length
* 80073 80172: gap of unknown length
* 80173 81415: contig of 1243 bp in length
* 81416 82576: gap of unknown length
* 82577 82576: contig of 1061 bp in length.

FEATURES

Source

Location/Qualifiers
1 82576

1. .82576

```
/organism="Rattus norvegicus"  
/db_xref="taxon:10116"
```

```
/db_xref="taxon:10116"
/clone="CH230-173010"
```

```
/clone="CH230-173010"
```

Query Match	7.68;	Score	40.6;	DB	2;	Length	82576;
Best Local Similarity	49.38;	Prod	No	0	00.		

Best Local Similarity	49.38;	Pred. No. 0.99;
Matches 106;	Conservative	0;
Mismatches	109;	Indels 0;
Gaps	0;	

Matches 106; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

95 AGAAACTCTGTTATGGATCACAACAACACACCATTTCAAATGGTTACCGCAGACACACA 154

Db 49949 AAACACATTATTAGGCATTCATGACCAATTTTATCTCAAAATGTGAAGAAACCATTGT 50008

155 GCCCAACACACAGGGTGTGGAGCCAATTCTCTGAACGTCACGATATGTACTCTTC 214

Db 50009 TTTTCACAACAATGGGCAGATAAAAAAGCAGCTGCAAGGGTTGGGGATTTCAGCTCAGTG 50068

215 GTACCACCCTTTACCTAAAAACATGCAGAGGAGATTCTACACCTAAGAAACAGAGAA 274

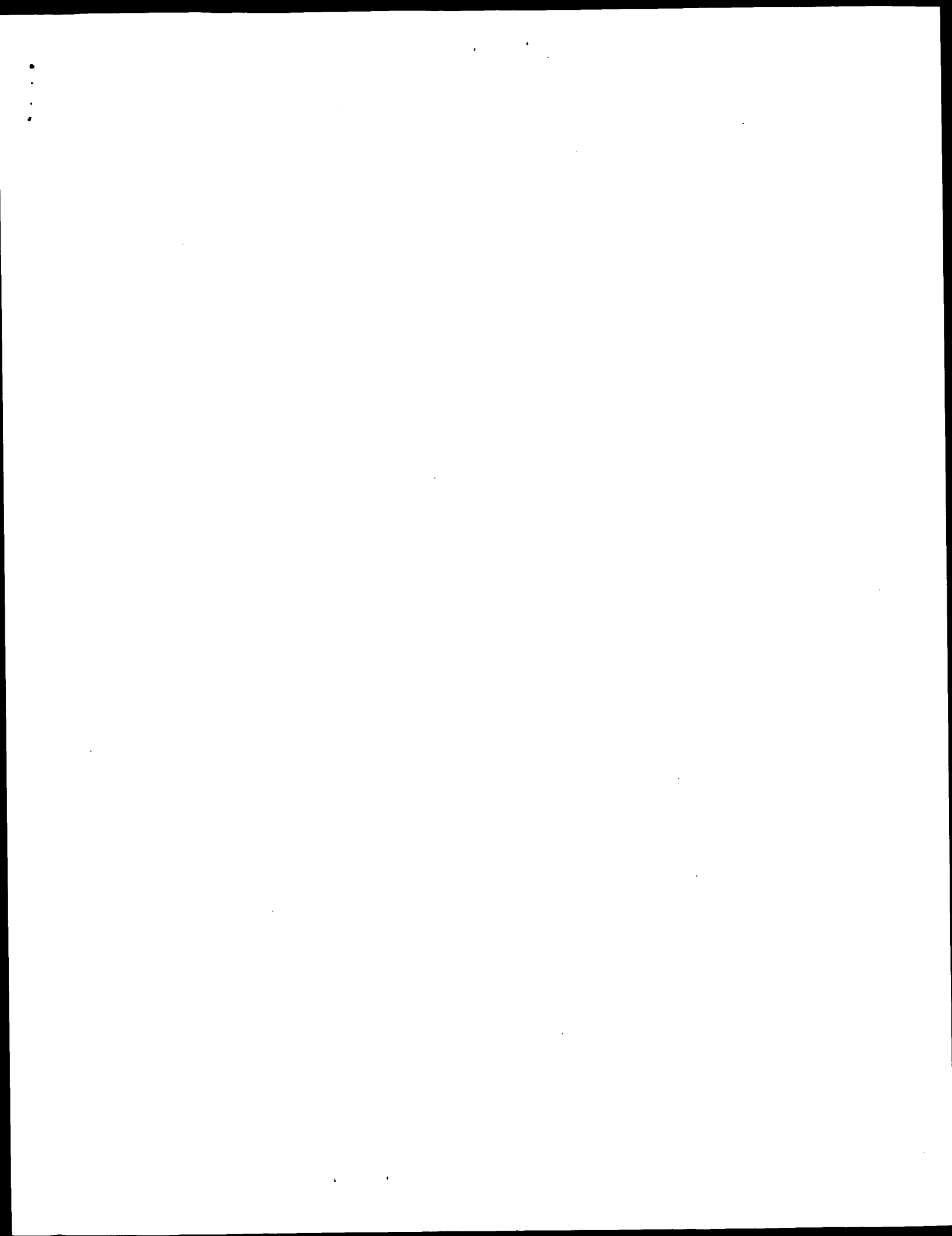
Db 50069 GTAGACGCTTGCCTAGCAAGCACAAGGCCCTGGATTCCGTCCAGCTCAAAAAAAAAA 50128

275 AACAAAGAAGAACCGCCTTTGGAGCCGGAAT 309

Db 50129 AAAAAAAAAAGCAGCTGCACCTTTT TAGACCTCAAT 50163

Search completed: October 19, 2002, 06:35:17
Job time : 2448 secs

Job time : 2448 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2002, 04:30:01 ; Search time 201 Seconds
(without alignments)
4544.271 Million cell updates/sec

Title: AI327498

Perfect score: 532
Sequence: 1 AATTCGGCAGTACAGACACCC.....CATTAATTCCTGCAATCAAT 532

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*

1:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.8	7.3	5688	23	ABL26543
2	38.8	7.3	7588	23	ABL26542
3	38	7.1	13631	23	ABL20354
4	37	7.0	400	23	AA575460
5	37	7.0	963	23	AA568580
6	35.8	6.7	12507	24	ABL32298
7	35.6	6.7	51	22	AA183398
8	35.2	6.6	364	22	AA184802
9	35.2	6.6	966	21	AA42768

10	35.2	6.6	1235	21	AA50838	Arabidopsis thaliana
11	35.2	6.6	1239	21	AA533292	Arabidopsis thaliana
12	34.2	6.4	496	22	AA191273	Human polynucleoti
13	34.2	6.4	2022	22	AA184529	E. coli growth and
14	33.8	6.4	663	23	AA565332	DNA encoding novel
15	33.8	6.4	663	23	AA580842	DNA encoding novel
16	33.8	6.4	793	22	AA195022	Human neuroblastom
17	33.8	6.4	5163	19	AAV20700	Cryptosporidium pa
18	33.8	6.4	5163	21	AAA61849	ORF encoding a por
19	33.8	6.3	5318	21	AAA61848	DNA encoding a por
20	33.6	6.3	1518	23	AA553583	Helicobacter pylori
21	33.6	6.3	5318	19	AAV20701	Cryptosporidium pa
22	33.6	6.3	7286	20	AA57224	WO 9923223 Seq ID
23	33.6	6.3	7938	20	AA57235	WO 9923223 Seq ID
24	33.4	6.3	466	22	AA182085	Human immune/haema
25	33.4	6.3	617	22	AAK71856	Human immune/haema
26	33.4	6.3	941	22	AAK71857	Human immune/haema
27	33.4	6.3	9516	22	AA542103	Genomic sequence #
28	33.2	6.2	5181	23	ABL10975	Arabidopsis thaliana
29	33.2	6.2	14974	23	ABL10974	Drosophila melanog
30	33	6.2	441	23	AA575453	DNA encoding novel
31	33	6.2	948	23	ABL03799	Drosophila melanog
32	33	6.2	2067	21	AA550488	soybean sulphate p
33	33	6.2	3399	17	AA705868	Chicken leucocyte
34	33	6.2	5813	24	ABL33804	Human immune syste
35	33	6.2	6151	24	ABL33611	Human immune syste
36	33	6.2	8305	24	ABL33569	Human immune syste
37	32.8	6.2	1686	16	AA087587	DNA encoding leuco
38	32.8	6.2	4590	22	AAH24065	Yeast AOD9604-asso
39	32.8	6.2	6071	24	AAH22316	Chemically treated
40	32.8	6.2	6853	24	ABL32311	Human immune syste
41	32.8	6.2	8446	24	ABL32311	Human immune syste
42	32.8	6.2	10877	21	AA555620	Human BMP-7 gene 5
43	32.6	6.1	5014	19	AAV58286	S. pyogenes SP-14-
44	32.4	6.1	357	22	AA187311	Human polynucleoti
45	32.4	6.1	1664	20	AA528072	Heparin sulphate 6

ALIGNMENTS

RESULT 1
ID ABL26543
ID ABL26543 standard; DNA: 5688 Bp.

XX ABL26543;
XX

DT 26-MAR-2002 (first entry)
XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 31102.
XX

KW Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaeaceutical; gene; ds.
XX

OS Drosophila melanogaster.
XX

PN WO2001/1042-A2.
XX

PD 27-SEP-2001.
XX

PF 23-MAR-2001; 2001WO-US09231.
XX

PR 23-MAR-2000; 2000US-191637P.
XX

PR 11-JUL-2000; 2000US-0614150.
XX

PA (PEKE) PE CORP NY.
XX

PL Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX

DR New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
PT

[illegible]

Query Match	7.3%:	Score 38.8:	DB 23:	Length 7588:
Best Local Similarity	54.1%:	Pred. NO. 0.39:	Mismatches 67:	Gaps 0:
Matches 79:	Conservative 0:	Mismatches 67:	Indels 0:	Gaps 0:
115	ACAACAACAACACCAATTCATAAATGGTTACCCGACGACACAGGCCAACAACACAGGGTGG	174		
Db	886	ACACACACAGACACATCTCCAAACGAGAAATTCACACACAGAAACATGCAAAACGACGA	827	
Qy	175	TGGAAGCCCAATTTCTGTAACGTCAGAGATATGTACTCTTCGTACCAACCTTTACCTAAAA	234	
Db	826	ATCAACAACAACAAACCAAGACATCCAAAGCAAGAAATCCAGAACAAACCAAGAACAGACA	767	
Qy	235	ACATGCAGAGAGGAGATTTCTACAAACCT	260	
Db	766	ATGAGCGAGAGAGAAACCAACCT	741	
RESULT 3				
ID	ABL20354	standard: DNA; 13631 BP.		
AC	ABL20354:			
XX	26-MAR-2002	(first entry)		
XX	Drosophila melanogaster	genomic polynucleotide	SEQ ID NO 12535.	
XX	Drosophila	developmental biology; cell signalling; insecticide;		
XX	pharmaceutical; gene; ds.			
XX	Drosophila melanogaster.			
XX	WO200171042-A2.			
XX	27-SEP-2001.			
XX	23-MAR-2001;	2001WO-US09231.		
XX	23-MAR-2000;	2000US-191637P.		
XX	11-JUL-2000;	2000US-0614150.		
XX	(PEKE)	PE CORP NY.		
XX	Venter JC,	Adams M, Li PWD,	Myers EW;	
XX	WPI;	2001-656860/75.		
XX	New isolated nucleic acid	detection reagent for detecting 1000 or more		
XX	genes from Drosophila	and for elucidating cell signalling and cell-cell		
XX	interactions -			
XX	Claim 1;	SEQ ID NO 12535; 21pp + Sequence Listing;	English.	
XX	The invention relates to an isolated nucleic acid	detection reagent		
XX	capable of detecting 1000 or more genes from	Drosophila. The invention is		
XX	useful in developmental biology and in elucidating	cell signalling and		
XX	cell-cell interactions in higher eukaryotes for	the development of		
XX	insecticides, therapeutics and pharmaceutical	drugs. The invention		
XX	discloses genomic DNA sequences (AB16176-AB130511),	expressed DNA		
XX	sequences (AB101840-AB16175) and the encoded	proteins		
XX	(ABB57737-ABB72072).			
XX	The sequence data for this patent did not form	part of the printed		
XX	specification, but was obtained in electronic	format directly from WIPO		
XX	at ftp.wipo.int/pub/published_pct_sequences.			
XX	Sequence 7588 BP; 1405 A; 1605 C; 2192 G;	2386 T; 0 other;		

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 13631 BP; 4097 A; 3233 C; 2518 G; 3783 T; 0 other;

Query Match 7.1%; Score 38; DB 23; Length 13631;
Best Local Similarity 50.5%; Pred. No. 0.9;
Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 112 ATCCAGACAGACACACCTTTAAATGGTTACCGACACAGACCCACACACACGAGG 171
DB 6532 ACCACACAGAC 6591
QY 172 TGGTGAAGCCAAATTCCTGAAGCTCCAGGATATGCTACTCTTGTGACACCTTAACTAA 231
DB 6592 ACAGAC 6651
QY 232 AAAACATGACAGACAGATTTCTACACCTTAAGAAACAGAAACAAAGAGAGACGGC 291
DB 6652 ACAACTACTACTACGACACACGACACACACACACACACACACACACACACACG 6711
QY 292 AC 293
DB 6712 AC 6713

RESULT 4

AA575460 standard; cDNA; 400 BP.

AA575460;

13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #11264.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG11273.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

Claim 1; SEQ ID No 11264; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensic, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 400 BP; 215 A; 34 C; 131 G; 19 T; 1 other;

Query Match 7.0%; Score 37; DB 23; Length 400;
Best Local Similarity 58.2%; Pred. No. 0.37;
Matches 64; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 231 AAAACATGACAGACAGATTTCTACACCTTAAGAAACAGAAACAAAGAGAGACGGC 290
DB 257 AAG 316
QY 291 CACTTTGAGCGCGAATGACACACCTGCGCCGCCAGGCTCCTACCTCCGCC 340
DB 317 AAGGAG 366

RESULT 5

AAS68580 standard; cDNA; 963 BP.

AAS68580;

13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #4384.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG04393.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

Claim 1; SEQ ID No 4384; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques

XX sequence 963 BP; 353 A; 171 C; 268 G; 169 T; 2 other

231	AAAAACATGCAGAGAGCATTTCTACAACCTAAGAAACAGAAAGAAAAACAAGAACCCG	239
QY		
257	TCTGACACAG	265

XX

XX	Human immune system associated gene	SEQ ID NO:	271
DE			

XX
XX
Homo sapiens

XX

XX
30-TTN-2000: 2000DE-1032529.

XX

XX

XX
XX
NO 371: 32np + Sequence Listing; German

XX
PS
XX

Query Match	Score	DB	Length
6.7%	35.8	24	1200

88 TGACCATAGAAACTCTGTTATGGATCACAACAACACCATTCAAAATGGTACCGA 14

[illegible]

..... 386

XX
XX
AAT.33398

Human SNP oligonucleotide #6606

XX

XX
XX
OF TTT - 2001

28-DEC-1999; 99US-0173419

XX
PT
shimkets RA, Leach

Polymorphic nucleic

XX The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amylolip proteins, angiotensin,
CC apolipoproteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukemia), diseases of the nervous system and an infection of pathogenic
CC organisms.

XX
SQ Sequence 51 BP; 15 A; 17 C; 11 G; 8 T; 0 other;

Query Match 6.7%; Score 35.6; DB 22; Length 51;
Best Local Similarity 82.0%; Pred. No. 0.39;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 140 TTACCGACGACACACACACACACAGGCTGAGGAGCAATTTCC 189
DB 2 TTACCGACGACACACACACACAGGCTGAGGAGCAATTTCC 51

RESULT 8
AA184802
ID AA184802 standard; cDNA; 364 BP.
XX
AC AA184802;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4862.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PE 26-FEB-2001; 2001WO-US04927.
XX
PF 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
XX
P-PSDB: AAO04871.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
PS Claim 1; SEQ ID NO 4862; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation and which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 364 BP; 172 A; 58 C; 71 G; 63 T; 0 other;

Query Match 6.6%; Score 35.2; DB 22; Length 364;
Best Local Similarity 53.7%; Pred. No. 1.2;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 151 CACAGCCCAACACACACAGGCTGAGGAGCAATTTCTGAGCTCCAGATATGACT 210
DB 1 CCCAGCGCTCCGCCACAGGCTCGATCGGCCCATATCCGAAATGTGTATACCT 60

QY 211 CTTCGTACCACTTTACTTAAACATGACAGAGAGATTTCTCAACCTAAGAACAGA 270
DB 61 TCCCGTACTTACCAAAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 120

QY 271 AGAAAAACAAAGAGAA 286
DB 121 AAAAAAAGAAAAAAGAA 136

RESULT 9
AAC42768
ID AAC42768 standard; DNA; 966 BP.
XX
AC AAC42768;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36776.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 23-APR-1999; 99US-0130891.
XX
PR 28-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 30-APR-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 21-JUN-1999; 99US-0139750.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143540.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 23-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147280.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

[illegible]

ID XX	AAC50838 standard; DNA; 1235 BP.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

AC AAC50838;
XX

18-OCT-2000 (first entry)

	Microopsis	DNA fragment SEQ ID NO:	66323
XX	thaliana		
XX			
Unbeidentf.			

protein identification; signal transduction pathway; metabolic pathway; protein expression control; genetic mapping; gene expression control;

metabolic pathway; promoter; termination sequence; ss.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999;	99US-0123180.
09-MAR-1999;	99US-0123548

23-MAR-1999; 99US-0125348.
25-MAR-1999; 99US-0125788.

23-MAR-1999;	99US-0126264.
29-MAR-1999;	99US-0126785.

R	01-APR-1999;	99US-0127462.
R	06-APR-1999;	99US-0127462.

00 APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.
16-APR-1999

R	16-APR-1999;	99US-0129845.
R	19-APR-1999;	99US-0130077

21-APR-1999; 99US-0130077.
23-APR-1999; 99US-0130449.

23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891

28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0131449.

30-APR-1999; 99US-0132048.
30-APR-1999; 99US-0132407.

04-MAY-1999;	99US-0132484.
05-MAY-1999.	99US-0132487.

05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132486.

06-MAY-1999; 99US-0132487.

1

100

PR	17-MAY-1999	9905-0132863
PR	11-MAY-1999	9905-0134256
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134219
PR	14-MAY-1999	9905-0134221
PR	14-MAY-1999	9905-0134768
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134768
PR	20-MAY-1999	9905-0134768
PR	21-MAY-1999	9905-0135124
PR	22-MAY-1999	9905-0135123
PR	25-MAY-1999	9905-0135629
PR	27-MAY-1999	9905-0136021
PR	28-MAY-1999	9905-0136392
PR	01-JUN-1999	9905-0136782
PR	03-JUN-1999	9905-0137222
PR	04-JUN-1999	9905-0137528
PR	07-JUN-1999	9905-0137502
PR	08-JUN-1999	9905-0137724
PR	10-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	14-JUN-1999	9905-0138847
PR	16-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139412
PR	17-JUN-1999	9905-0139453
PR	18-JUN-1999	9905-0139452
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139464
PR	18-JUN-1999	9905-0139465
PR	18-JUN-1999	9905-0139466
PR	18-JUN-1999	9905-0139467
PR	18-JUN-1999	9905-0139468
PR	18-JUN-1999	9905-0139469
PR	18-JUN-1999	9905-0139470
PR	18-JUN-1999	9905-0139471
PR	18-JUN-1999	9905-0139472
PR	18-JUN-1999	9905-0139473
PR	18-JUN-1999	9905-0139474
PR	18-JUN-1999	9905-0139475
PR	18-JUN-1999	9905-0139476
PR	18-JUN-1999	9905-0139477
PR	18-JUN-1999	9905-0139478
PR	18-JUN-1999	9905-0139479
PR	18-JUN-1999	9905-0139480
PR	18-JUN-1999	9905-0139481
PR	18-JUN-1999	9905-0139482
PR	18-JUN-1999	9905-0139483
PR	18-JUN-1999	9905-0139484
PR	18-JUN-1999	9905-0139485
PR	18-JUN-1999	9905-0139486
PR	18-JUN-1999	9905-0139487
PR	18-JUN-1999	9905-0139488
PR	18-JUN-1999	9905-0139489
PR	18-JUN-1999	9905-0139490
PR	18-JUN-1999	9905-0139491
PR	18-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139493
PR	18-JUN-1999	9905-0139494
PR	18-JUN-1999	9905-0139495
PR	18-JUN-1999	9905-0139496
PR	18-JUN-1999	9905-0139497
PR	18-JUN-1999	9905-0139498
PR	18-JUN-1999	9905-0139499
PR	18-JUN-1999	9905-0139500
PR	18-JUN-1999	9905-0139501
PR	18-JUN-1999	9905-0139502
PR	18-JUN-1999	9905-0139503
PR	18-JUN-1999	9905-0139504
PR	18-JUN-1999	9905-0139505
PR	18-JUN-1999	9905-0139506
PR	18-JUN-1999	9905-0139507
PR	18-JUN-1999	9905-0139508
PR	18-JUN-1999	9905-0139509
PR	18-JUN-1999	9905-0139510
PR	18-JUN-1999	9905-0139511
PR	18-JUN-1999	9905-0139512
PR	18-JUN-1999	9905-0139513
PR	18-JUN-1999	9905-0139514
PR	18-JUN-1999	9905-0139515
PR	18-JUN-1999	9905-0139516
PR	18-JUN-1999	9905-0139517
PR	18-JUN-1999	9905-0139518
PR	18-JUN-1999	9905-0139519
PR	18-JUN-1999	9905-0139520
PR	18-JUN-1999	9905-0139521
PR	18-JUN-1999	9905-0139522
PR	18-JUN-1999	9905-0139523
PR	18-JUN-1999	9905-0139524
PR	18-JUN-1999	9905-0139525
PR	18-JUN-1999	9905-0139526
PR	18-JUN-1999	9905-0139527
PR	18-JUN-1999	9905-013952

PR	23-JUL-1999.	9905-0145218.
PR	23-JUL-1999.	9905-0145218.
PR	23-JUL-1999.	9905-0145224.
PR	26-JUL-1999.	9905-0145376.
PR	27-JUL-1999.	9905-0145513.
PR	27-JUL-1999.	9905-0145513.
PR	27-JUL-1999.	9905-0145518.
PR	27-JUL-1999.	9905-0145591.
PR	28-JUL-1999.	9905-0145591.
PR	02-AUG-1999.	9905-0146386.
PR	02-AUG-1999.	9905-0146388.
PR	02-AUG-1999.	9905-0146389.
PR	03-AUG-1999.	9905-0147204.
PR	04-AUG-1999.	9905-0147302.
PR	05-AUG-1999.	9905-0147302.
PR	05-AUG-1999.	9905-0147260.
PR	05-AUG-1999.	9905-0147303.
PR	06-AUG-1999.	9905-0147416.
PR	09-AUG-1999.	9905-0147493.
PR	09-AUG-1999.	9905-0147935.
PR	11-AUG-1999.	9905-0148171.
PR	11-AUG-1999.	9905-0148319.
PR	13-AUG-1999.	9905-0148565.
PR	13-AUG-1999.	9905-0148684.
PR	13-AUG-1999.	9905-0148684.
PR	16-AUG-1999.	9905-0149368.
PR	17-AUG-1999.	9905-0149175.
PR	18-AUG-1999.	9905-0149426.
PR	20-AUG-1999.	9905-0149723.
PR	20-AUG-1999.	9905-0149929.
PR	23-AUG-1999.	9905-0149929.
PR	23-AUG-1999.	9905-0149930.
PR	25-AUG-1999.	9905-0150566.
PR	26-AUG-1999.	9905-0150884.
PR	27-AUG-1999.	9905-0151065.
PR	27-AUG-1999.	9905-0151066.
PR	30-AUG-1999.	9905-0151068.
PR	31-AUG-1999.	9905-0151303.
PR	01-SEP-1999.	9905-0151438.
PR	01-SEP-1999.	9905-0151930.
PR	07-SEP-1999.	9905-0153363.
PR	10-SEP-1999.	9905-0153070.
PR	13-SEP-1999.	9905-0153758.
PR	15-SEP-1999.	9905-0154018.
PR	16-SEP-1999.	9905-0154039.
PR	20-SEP-1999.	9905-0154779.
PR	23-SEP-1999.	9905-0155139.
PR	24-SEP-1999.	9905-0155486.
PR	28-SEP-1999.	9905-0155659.
PR	28-SEP-1999.	9905-0156458.
PR	28-SEP-1999.	9905-0156596.
PR	04-OCT-1999.	9905-0157117.
PR	04-OCT-1999.	9905-0157753.
PR	05-OCT-1999.	9905-0157665.
PR	06-OCT-1999.	9905-0158029.
PR	07-OCT-1999.	9905-0158322.
PR	08-OCT-1999.	9905-0158323.
PR	12-OCT-1999.	9905-0159253.
PR	13-OCT-1999.	9905-0159254.
PR	13-OCT-1999.	9905-0159295.
PR	13-OCT-1999.	9905-0159329.
PR	14-OCT-1999.	9905-0159330.
PR	14-OCT-1999.	9905-0159330.
PR	14-OCT-1999.	9905-0159331.
PR	14-OCT-1999.	9905-0159637.
PR	14-OCT-1999.	9905-0159584.
PR	14-OCT-1999.	9905-0159584.
PR	18-OCT-1999.	9905-0160741.
PR	18-OCT-1999.	9905-0160741.
PR	21-OCT-1999.	9905-0160767.
PR	21-OCT-1999.	9905-0160768.
PR	21-OCT-1999.	9905-0160770.
PR	21-OCT-1999.	9905-0160814.
PR	21-OCT-1999.	9905-0160815.

[illegible]

AAC332921

XX DT	17-OCT-2000 (first entry)
17-OCT-2000	(first entry)

	DNA fragment SEQ ID NO:
XX	2492

DE Arabidopsis thaliana; gene expression control; XX

Hybridisation assay; genetic mapping; signal transduction pathway; protein identification; signal transduction sequence; ss

metabolic pathway; P₁

Arabidopsis t

XX
PN EP1033405-A2

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825
PR

PR	05-MAR-1999;	99US-0123180
PR	09-MAR-1999;	99US-0123548

PR 23-MAR-1999; 99US-0125788
PR 23-MAR-1999; 99US-0126264

PR	23-MAR-1999;	99US-0126785
PR	29-MAR-1999;	99US-0127462

PR	01-APR-1999;	99US-0128234
PR	06-APR-1999;	99US-0128714

PR	08-APR-1999;	99US-0120714
PR	16-APR-1999;	99US-0129845

PR 19-APR-1999; 99US-0130077
PR 31-APR-1999; 99US-0130449

PR	23-APR-1999;	99US-0130510
PR	23-APR-1999;	99US-0130891

PR	23-APR-1999;	99US-0131449
PR	28-APR-1999;	99US-0132048

PR 30-APR-1999; 99US-0132407
PR 30-APR-1999; 99US-0132486

PR 04-MAY-1999; 99US-0132480
PR 05-MAY-1999; 99US-0132485

PR 06-MAY-1999; 9905-013248

```
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0133256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
```

XX	
CC	The invention relates to human polynucleotides (AA179941-AA193841) and
CC	the encoded proteins (AA000010-AA013910) that exhibit activity elating to

CC (1) and the purification of the proteins, the purified proteins can be
CC used to generate reagents and screen small molecule libraries or other
CC candidate compound libraries for compounds that can be further developed

to yield novel antimicrobial compounds. In addition, nucleic acid probes complementary to (I) that are specific for particular species of microorganisms can be used to identify particular microorganism species in clinical specimens, therefore, providing a rapid and dependable method by which to identify the causative agents of a bacterial infection. Also, antibodies generated against proteins translated from mRNA transcribed from proliferation-requiring sequences can also be used to screen for specific microorganisms that produce such proteins in a species-specific manner. AAH84371 and AAH84670 represent sequencing primers used in the isolation of *E. coli* growth and proliferation related sequence, which are used in an example from the present invention.

Sequence 2022 BP; 495 A; 488 C; 593 G; 446 T; 0 other;

Query Match 6.4%; Score 34.2; DB 22; Length 2022;
Best Local Similarity 55.5%; Pred. No. 5.4;
Matches 66; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 334 CTGCCCCCTTCTAAGAACTCGAATGCGGTCTGCTGCTTGCCTTACGAGGCTCT 393
DB 312 CTCTCTGCTTTGTAAGATACGATCAGCTTGTCTTAAAGATTGACCGAGGCT 371
QY 394 GATTAAACCCAGAGGAGTTCTCTCCGAGAGAAAGTGAGCTAGCTTACCGAGAGA 452
DB 372 GATTGAAGATGACAAAGTTCTCTCGAACAACGATTGACCATCTTAACGAGAGA 430

RESULT 14

AA565332 standard; cDNA; 663 BP.

AA565332;

13-FEB-2002 (first entry).

DNA encoding novel human diagnostic protein #1136.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dymanac RT, Liu C, Tang YT;

WPI: 2001-639362/73.

P-PSDB; ABG01145.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID No 1136; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving

(II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AA564197-AA594564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 663 BP; 240 A; 95 C; 116 G; 212 T; 0 other;

Query Match 6.4%; Score 33.8; DB 23; Length 663;
Best Local Similarity 45.6%; Pred. No. 4.3;
Matches 119; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 28 CAGAGAAATGTCAGTGGTTTGCTTACAGCAAAAGTCACTTGTCCAGGCCCGAGTGG 87
DB 355 CATCAAGATTAAAGTATTTGCTTACGCTAAATGTCATTAATTTAACTGAG 414
QY 88 TGACCATGAAACACTCTGTTATGATCACAACACACCATTCAAATGGTTACCGCA 147
DB 415 TTATGAAAAAAGTATTTGGCAGAAAAAATTAAGAGATGATATGTCAGACA 474
QY 148 GCACACAGCCCAACACACAGGTTGTGAGAACCAATTTCTGACGTCACAGATATGT 207
DB 475 GCCACAGGGGAAAGAGAACAGTGTGAAGAACCATTAATTTGAAACATGTAGACAC 534
QY 208 ACTCTTGTACACCTTTTACCTAAACATGCAAGAGAGATTTCTAACACCTAAGAAC 267
DB 535 ATGAGAGAAATTAATCTTAACCAAGAACGAAAGGGAAGAGGTGTTCAATTCTAATA 594
QY 268 AGAAGAAAAACAAGAGAAC 288
DB 595 GGAAGAAAAAGATTAATATATC 615

RESULT 15

AA580842 standard; cDNA; 663 BP.

AA580842;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #16646.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dymanac RT, Liu C, Tang YT;

WPI: 2001-639362/73.

P-PSDB; ABG16655.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX
PS Claim 1: SEQ ID No 16646; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPRO
CC at ftp.wipro.int/pub/published_pct_sequences.

XX
SQ Sequence 663 BP; 240 A; 95 C; 116 G; 212 T; 0 other;

Query Match 6.4%; Score 33.8; DB 23; Length 663;
Best Local Similarity 45.6%; Pred. No. 4.3;

Matches 119; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 28 CAGAGATGTCAGGTTTGCTTACAGACAAAGTCACCTTGTCCAGGCCCAAGTGG 87
DB 355 CATTCAAAGATTAAAGTATTGCTTTCATTAAGCTAAATGTCATATTAATTTAACTGAG 414
QY 88 TGACCATAGAAAACCTGTTATGATCACAACAAACACACATTCMAAATGGTTACCGCA 147
DB 415 TTATTGAAAAAAAGTATTGGCGAAGAAAAAATTAAAGAGATGTATATATGTCAAGACA 474
QY 148 GCACACAGCCCAACACACAGAGGTGTGGAAGCCAAATTCCTGACGTCACGATATGT 207
DB 475 GCCACAGAGGGAAGAGAACAGTGTGAAGAACCATATATGATTGGAAACATGTAGAAC 534
QY 208 ACTCTTGCTACCACTTACTTAATAAAACATGCAGAGAGATTTCTACACCTTAAGAAC 267
DB 535 ATGAGGAAATTTACTTAAACCAAGAAAGCAAAAGGGGAAAGGTGTCTATTTCTAAAAA 594
QY 268 AGAAGAAAACAAGAAAGAAC 288
DB 595 GGAAGAAAAGATTAATATC 615

Search completed: October 19, 2002, 05:54:31
Job time : 225 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2002, 05:44:31 ; Search time 44 Seconds

(without alignments)
2969,932 Million cell updates/sec

Title: AI327498

Sequence: 1 AATTCGGCAGTACGACACCC.....CATTAATTCCTGTCATCAT

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_NA:*
2: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
3: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
5: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
6: /cgn2_6/prodata/2/ina/PCUTUS.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	38.8	7.3	7218	1	US-08-232-463-14
2	33.8	6.4	5163	3	US-08-700-651-1
3	33.8	6.4	5163	3	US-08-928-361B-4
4	33.8	6.4	5318	3	US-08-700-651-2
5	33.8	6.4	5318	3	US-08-928-361B-3
6	33.6	6.3	7286	3	US-09-331-581-3
7	33.6	6.3	7938	3	US-09-331-581-14
8	33.2	6.2	289	4	US-09-007-005-17
9	33.2	6.2	289	4	US-09-444-796-17
10	32.6	6.1	5014	4	US-09-381-862-3
11	32	6.0	2139	3	US-09-026-343-1
12	32	6.0	6924	1	US-08-015-973-2
13	32	6.0	6924	2	US-08-448-164-2
14	31.8	6.0	1099	4	US-08-944-604-19
15	31.8	6.0	1262	4	US-08-944-604-17
16	31.8	6.0	1331	4	US-08-944-604-17
17	31.8	6.0	4621	1	US-08-770-761A-1
18	31.8	6.0	5511	3	US-08-928-361B-2
19	31.8	6.0	7334	3	US-08-928-361B-1
20	31.8	6.0	19124	2	US-08-487-828B-13
21	31.2	5.9	198	5	PCT-US95-10668-1
22	30.8	5.8	2657	2	US-08-974-565C-8
23	30.8	5.8	2657	2	US-09-255-748-8
24	30.6	5.8	1296	4	PCT-US91-00909-5
25	30.6	5.8	5024	4	US-09-307-143-1
26	30.6	5.8	246240	2	US-08-724-394A-20
27	30.6	5.8	246240	2	US-08-724-394A-20

C 28	30.6	5.8	246240	2	US-08-724-394A-21	Sequence 21, Appl
C 29	30.6	5.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 30	30.2	5.7	1066	1	US-08-157-101A-4	Sequence 4, Appl
C 31	30.2	5.7	1398	2	US-08-896-320-2	Sequence 2, Appl
C 32	30.2	5.7	6601	3	US-09-356-952-10	Sequence 10, Appl
C 33	29.8	5.6	4808	2	US-08-231-193A-10	Sequence 10, Appl
C 34	29.8	5.6	4808	2	US-08-486-273A-10	Sequence 10, Appl
C 35	29.8	5.6	4808	3	US-08-940-035A-10	Sequence 10, Appl
C 36	29.8	5.6	4808	4	US-08-436-322B-9	Sequence 9, Appl
C 37	29.8	5.6	198	5	PCT-US95-10668-3	Sequence 9, Appl
C 38	29.6	5.6	198	5	PCT-US95-10668-4	Sequence 9, Appl
C 39	29.6	5.6	1107	2	US-08-991-300-1	Sequence 1, Appl
C 40	29.6	5.6	1219	4	US-09-230-380-10	Sequence 1, Appl
C 41	29.6	5.6	1325	1	US-08-306-691B-51	Sequence 10, Appl
C 42	29.6	5.6	2151	4	US-09-626-589-4	Sequence 51, Appl
C 43	29.6	5.6	2284	2	US-08-467-822-28	Sequence 4, Appl
C 44	29.6	5.6	2284	2	US-08-467-822-28	Sequence 28, Appl
C 45	29.6	5.6	2284	4	US-08-432-697-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c

/ Sequence 14, Application US/08232463

/ Patent No. 5670367

/ GENERAL INFORMATION:

/ APPLICANT: DORNER, F.

/ APPLICANT: SCHLEIFINGER, F.

/ APPLICANT: FALKNER, F. G.

/ TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

/ NUMBER OF SEQUENCES: 52

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Foley & Lardner

/ STREET: 1800 Diagonal Road, Suite 500

/ CITY: Alexandria

/ STATE: VA

/ COUNTRY: USA

/ ZIP: 22313-0299

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/232,463

/ FILING DATE:

/ CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US/07/935,313

/ FILING DATE: 26-AUG-1991

/ ATTORNEY/AGENT INFORMATION:

/ NAME: BENT, Stephen A.

/ REGISTRATION NUMBER: 29,768

/ REFERENCE/DOCKET NUMBER: 30472/114

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (703)836-9300

/ TELEFAX: (703)683-4109

/ TELETYPE: 899149

/ INFORMATION FOR SEQ ID NO: 14:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 7218 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ IMMEDIATE SOURCE:

/ CLONE: pTZ9pC-F15

/ US-08-232-463-14

Query Match

7.3%; Score 38.8; DB 1; Length 7218;

Best Local Similarity 3.4%; Pred. No. 0.083; Indels 0; Gaps 0;
Matches 13; Conservative 208; Mismatches 165;

```

QY 93 ATGAGAACTCTGTTATGATCACAACAACACCATTTCAAAATGTTACCGAGGACA 152
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1451 ATGAGAACTCTGTTATGATCACAACAACACCATTTCAAAATGTTACCGAGGACA 1392
QY 153 CAGCCACAACAACACGAGGTGTCAGGACATTTCTGAGCTCAGATATGACTCT 212
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1391 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1332
QY 213 TCGACACCTTACTTAAATAATGACAGAGAGATTTCTCAACCTTAAGAAAGAG 272
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1331 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1272
QY 273 AAAACAAGAAGAACCGGCTTTGGAGCGGGAATGACACACGCGCCAGGCTCAC 332
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1271 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1212
QY 333 CCGGCCCTTGTCTAGGAATCTGGAATGGGTCTCTGCTGCTCAGGAGTCTC 392
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1211 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1152
QY 393 TGATTACCCAGAGGAGTTCTCTCCGAGAGAAGAGTGAAGTACCTAGCCAGAGA 452
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1151 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092
QY 453 GAAGGAGTGTGGTGAGAGGAGAAA 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1091 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1066

```

RESULT 2
US-08-700-651-1
Sequence 1, Application US/08700651B

```

; GENERAL INFORMATION:
; PATENT NO. 6015882
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; FILE REFERENCE: 480.19-4(HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 08/415,751
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5163
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-08-700-651-1

```

Query Match 6.4%; Score 33.8; DB 3; Length 5163;
Best Local Similarity 49.7%; Pred. No. 2.1;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

QY 115 ACAACAACAACACATTTCAAAATGTTACCGAGCAGACAGCCCAACACAGAGGTGG 174
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 674 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 733
QY 175 TGAAGCCATTTCTGAGAGTCGAGATATGTAATCTTCTGACACCTTTACTTAAAA 234
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 734 ACAGTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 793
QY 235 ACATGCAAGAGAGATTTTACACCTTACAGAAACAAGAAAAAAGAAAGAAC 287
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 794 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 846

```

RESULT 3
US-08-928-361B-4
Sequence 4, Application US/08928361B

```

; GENERAL INFORMATION:
; PATENT NO. 6071518
; APPLICANT: Petersen, Carolyn
; APPLICANT: PETERSEN, CAROLYN
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5163 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-4

```

Query Match 6.4%; Score 33.8; DB 3; Length 5163;
Best Local Similarity 49.7%; Pred. No. 2.1;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

QY 115 ACAACAACAACACATTTCAAAATGTTACCGAGCAGACAGCCCAACACAGAGGTGG 174
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 673 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 732
QY 175 TGAAGCCATTTCTGAGAGTCGAGATATGTAATCTTCTGACACCTTTACTTAAAA 234
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 733 ACAGTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 792
QY 235 ACATGCAAGAGAGATTTTACACCTTACAGAAACAAGAAAAAAGAAAGAAC 287
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 793 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 845

```

RESULT 4
US-08-700-651-2
Sequence 2, Application US/08700651B

```

; GENERAL INFORMATION:
; PATENT NO. 6015882
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI

```

Patent No. 6130070
GENERAL INFORMATION:
APPLICANT: TOHDA, Hideki
APPLICANT: HAMA, Yuko
APPLICANT: KUMAGAI, Hiromichi
TITLE OF INVENTION: INDUCTION PROMOTER GENE AND SECRETORY SIGNAL GENE

Oy	239	GCGAGGAGATCTTCAACACTAAGAAGCAGAAAAACAAGAGAACGGCATTTTG	298
		: : : : : : : : : : : : : : : : : :	
Dd	6	GACRCRRARURRARCRRPARURRUBRCARARURRBRACRRARRURRRNRNSNR	65
Oy	299	GAGCCGGAATGCACACTGGCCCCAGGGCTCACCCCTGCCCTTGTCTAAGAAATCTCG	356
		: :	
Dd	66	NRSNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRR	122
Oy	359	AATGTGGCTTCCTTGCTGGCTTCACGCGAGTCTGATTAAACCAGAGGAGTTCCTC	418
Dd	126	NRSNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRNRR	182
Oy	419	CGACAGAAGAGTGGAGCTTAGCTTAAGCCAGAAAGAGAGATGTGGGTAGAGAGAAA	476
		: :	
Dd	186	NRSNRRNSRNRRNSRNRRNSRNRRNSRNRRNSRCRARGCRURRCGRUPARARCRUR	242
Oy	479	GACTTCGACGAGCGTGCAGAAAGAAATCA	506
		: : : : : : : : : : : : : : : : : : : : : :	
Dd	246	CRURRGRCRCRURAAAAAAAAAAAAAAAA	273

```

RESULT 10
US-09-381-862-3
: Sequence 3, Application US/09381862
: Patent No. 6245906
: GENERAL INFORMATION:
: APPLICANT: Ueyama, Hiroshi
: APPLICANT: Abe, Kanako
: APPLICANT: Keshi, Hiroyuki
: APPLICANT: Matsuhisa, Akio
: TITLE OF INVENTION: PROBES FOR THE DIAGNOSIS OF INFECTIONS
: TITLE OF INVENTION: CAUSED BY STREPTOCOCCUS PYOGENES
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive/6300 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/381,862
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 1997-71077
: FILING DATE: 25-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP98/01288
: FILING DATE: 23-MAR-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Cawley, Jr., Thomas A.
: REGISTRATION NUMBER: 40,944
: REFERENCE/DOCKET NUMBER: 19036/36274
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5014 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Streptococcus pyogenes
: STRAIN: Clinical Isolate SP-14-1
US-09-381-862-3

Query Match 6.1%; Score 32.6; DB 4; Length 5014;
Best Local Similarity 58.9%; Pred No. 4.7;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 78 GCCACGCTGGTACCATGAAAACCTCTGTTATGATCAACAACAACACATTCAAAAT 137
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3815 GCGAAGGTATTAAAGTGTTCCTAAATTATTCATGACACCAAAAACATATATAAAT 3874
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 GGTATCCGACGACACACCAACAACAACACACGCGT 172
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3875 GGTATGCTTGATGATAGACCAACAAGACGACATGT 3909
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-026-343-1
: Sequence 1, Application US/09026343
: Patent No. 6008018
: GENERAL INFORMATION:
: APPLICANT: Duan, D. ROYANNE
: APPLICANT: SHLATITERAD, ALI

```

```

?
? APPLICANT: CONAWAY, JOAN W.
? APPLICANT: CONAWAY, RONALD C.
? TITLE OF INVENTION: ELI2, A New Member of an ELI Family of
? TITLE OF INVENTION: RNA Polymerase II Elongation Factors
? NUMBER OF SEQUENCES: 34
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
? STREET: 1100 NEW YORK AVENUE, SUITE 600
? CITY: WASHINGTON
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3934
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/026,343
? FILING DATE: Herewith
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/038,447
? FILING DATE: 19-FEB-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: GOLDSTEIN, JORGE A.
? REGISTRATION NUMBER: 29,021
? REFERENCE/DOCKET NUMBER: 1488.0880001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2600
? TELEFAX: (202) 372-2540
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2139 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: both
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 94..2013
?
? US-09-026-343-1
?
? Query Match 6.0%; Score 32; DB 3; Length 2139;
? Best Local Similarity 58.3%; Pred. No. 4.9;
? Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0
?
? QY 248 GATTCTCAACCTTAAGAAAGAGAAAGAAACAAGAAAGACCGGCACCTTTGGAGCCGGAA 307
? ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
? DB 1100 GAATTCCTCACTGACGACAGACAGATGACCACTAAATGTCATTGTAATCCACCA 1159
?
? QY 308 ATGACACACTGGCCCCAGGGCTTCACCCCTGCGCCCTTG 343
? ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
? DB 1160 GTGAAATATCGCTGCAGGCTCCCACTGCCCCCTG 1195
?
? RESULT 12
? US-08-015-973-2
? Sequence 2, Application US/08015973
? Patent No. 5604094
? GENERAL INFORMATION:
? APPLICANT: Schliesinger, Joseph
? TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
? TITLE OF INVENTION: PHOSPHATASE-BETA
? NUMBER OF SEQUENCES: 9
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PENNIE & EDMONDS
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 10036-2711
?
? COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/015,973
FILING DATE: 10-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISTOCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6924
US-08-015-973-2

Query Match
Best Local Similarity 6.0%; Score 32; DB 1; Length 6924;
Matches 47; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 16 CACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAGTCACTTGTCCA 75
DB 710 CTCCTCCCTGCACAGACAGTGTGATGATGTTTAAAGATACAGTTAGCATCTCTG 769

QY 76 AGGCCCAGGTGG 87
DB 770 AAGCCAGTTGG 781

RESULT 13
US-08-448-164-2
Sequence 2, Application US/08448164
Patent No. 5925536
GENERAL INFORMATION:
APPLICANT: Schlensing, Joseph
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,164
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/015,973
FILING DATE: 10-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: MISTOCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-021

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6924
US-08-448-164-2

Query Match
Best Local Similarity 6.0%; Score 32; DB 2; Length 6924;
Matches 47; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 16 CACCCCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAGTCACTTGTCCA 75
DB 710 CTCCTCCCTGCACAGACAGTGTGATGATGTTTAAAGATACAGTTAGCATCTCTG 769

QY 76 AGGCCCAGGTGG 87
DB 770 AAGCCAGTTGG 781

RESULT 14
US-08-944-604-19
Sequence 19, Application US/08944604
Patent No. 6218131
GENERAL INFORMATION:
APPLICANT: KERSEY, SUSAN
APPLICANT: OBAR, ROBERT
TITLE OF INVENTION: MATERIALS AND METHODS FOR DETECTION OF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,604
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MEYERS, THOMAS C
REGISTRATION NUMBER: 36,989
REFERENCE/DOCKET NUMBER: MTP-021 (8395/24)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1099 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 61..948

OTHER INFORMATION: /product="BC-8 (Isoform B)"
US-08-944-604-19

Query Match 6.0%; Score 31.8; DB 4; Length 1099;
Best Local Similarity 49.7%; Pred. No. 4.3;
Matches 81; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 327 GCTACCCCTGCCCCCTTGTCTAAGAAATCCTGGAAATGTGGTCTCTCTGCTTCCCTTCACGG 386
DB 470 GCTCAAAATCCTCAGGTGTACATGACATCAAGATTGGGAACAGCCGGCTGGCCGATCC 529
QY 387 AGTCTGTATTAAACCCAGAGGAGTTCCTCCGAGAGAAAGTAGAGTACGTTAGCC 446
DB 530 AGATGCTCTGCTGCTTGTGATGTGTGCTCCATGACAGCAGAAATTTCCGCTGTGCA 589
QY 447 AGAAGAGAAAGGAGATGTGGGTGAGAGGAAAGACCTTCGACGG 489
DB 590 CTCATGAAAGGGCTTGGCTTTAAGGAAAGCAGCTTCACCG 632

RESULT 15

US-08-989-386-4
Sequence 4, Application US/08989386
Patent No. 5989860

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN ISOMERASE HOMOLOGS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,386

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0443 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1262 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: TMLRDT01

CLONE: 292808

US-08-989-386-4

Query Match 6.0%; Score 31.8; DB 2; Length 1262;
Best Local Similarity 49.7%; Pred. No. 4.6;
Matches 81; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 327 GCTACCCCTGCCCCCTTGTCTAAGAAATCCTGGAAATGTGGTCTCTCTGCTTCCCTTCACGG 386
DB 468 GCTCAAAATCCTCAGGTGTACATGACATCAAGATTGGGAACAGCCGGCTGGCCGATCC 527
QY 387 AGTCTGTATTAAACCCAGAGGAGTTCCTCCGAGAGAAAGTAGAGTACGTTAGCC 446
DB 528 AGATGCTCTGCTGCTTGTGATGTGTGCTCCATGACAGCAGAAATTTCCGCTGTGCA 587
QY 447 AGAAGAGAAAGGAGATGTGGGTGAGAGGAAAGACCTTCGACGG 489
DB 588 CTCATGAAAGGGCTTGGCTTTAAGGAAAGCAGCTTCACCG 630

Search completed: October 19, 2002, 06:54:48
Job time : 65 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 19, 2002, 04:58:21 ; Search time 1637 Seconds

(Without alignments)
4386.304 Million cell updates/sec

Title: AI327498

Perfect score: 532
Sequence: 1 AATTCGCACCTAGACACCC.....CATTAATTCCTGCAATCAT 532

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	532	100.0	532	9	AI327498
2	516.4	97.1	704	10	BG865668
3	516.4	97.1	742	10	BG871142
4	516.4	97.1	757	10	BG866610
5	516.4	97.1	1361	11	BC018582
6	514.8	96.8	874	10	BG974340
7	513.8	96.6	841	10	BG867642
8	511.6	96.2	806	10	BF299645
9	505.4	95.0	822	10	BF300481
10	504.4	94.8	740	10	BG866897
11	504.4	94.8	849	10	BF537367
12	504.4	94.8	891	10	BF535022
13	503.8	94.7	743	10	BG873314
14	502.8	94.5	988	10	BF300873
15	496.4	93.3	741	10	BG871007
16	494.8	93.0	708	10	BG868576
17	493.4	92.7	827	10	BG865197

18	487.4	91.6	675	9	AV249432
19	487.4	91.6	680	10	BF300518
20	485.4	91.2	860	10	BG866108
21	483.4	90.9	682	10	BF535168
22	482.4	90.7	729	10	BG870130
23	480.8	90.3	631	10	BF301693
24	480.4	90.3	727	10	BF539392
25	479	90.0	791	10	BG869048
26	478	89.8	810	10	BF301149
27	474.4	89.2	730	10	BG873295
28	470.2	88.4	740	10	BG869019
29	469.4	88.2	678	10	BG869045
30	468.6	88.1	874	10	BG870986
31	465.4	87.5	955	10	BF538505
32	464.4	87.3	718	10	BF540048
33	462.8	87.0	826	10	BG865309
34	462	86.8	851	10	BG872096
35	461.4	86.7	799	10	BG871270
36	459.8	86.4	703	10	BF534356
37	459.6	86.4	692	10	BG868162
38	459.4	86.4	692	10	BG870992
39	456.8	85.9	834	10	BF539409
40	456.6	85.8	672	10	BF538109
41	456.4	85.8	665	10	BF534301
42	453.8	85.3	734	10	BF534781
43	452.4	85.0	692	10	BF536890
44	449.8	84.5	671	10	BF534781
45	448.2	84.2	670	10	BF300142

ALIGNMENTS

RESULT 1
LOCUS AI327498
DEFINITION mb20a07.yl Soares mouse p3NM19.5 Mus musculus CDNA clone
IMAGE:329940 5' similar to SW:CAH6_SHEEP P08060 CARBONIC ANHYDRASE
VI : mRNA sequence.

ACCESSION AI327498.1 GI:4061927
VERSION AI327498
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 532)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

The WashU-HMT Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project

WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through INL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:211340

This read is a RESSEQUENCE of a previously sequenced mouse clone
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 476.
Location/Qualifiers

FEATURES

1..532
/organism="Mus musculus"
/db_xref="taxon:10090"

mRNA sequence.
 ACCESSION BG871142
 VERSION BG871142.1 GI:14221682
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLAM10840 row: e column: 13
 High quality sequence stop: 741.
 Location/Qualifiers
 1..742
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="4921716"
 /clone_lib="NCL_CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-Sport6; Site: 1; Motif: Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCL_CGAP Library."
 BASE COUNT 231 a 199 c 166 g 146 t
 ORIGIN
 Query Match 97.1%; Score 516.4; DB 10; Length 742;
 Best Local Similarity 99.8%; Pred. No. 3e-105;
 Matches 517; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 15 ACACCCCCCGCAGAGATGTCAGTGTGTGCTTAAAGACAAATCTTGTCC 74
 |||||||
 Db 107 ACACCCCCCGCAGAGATGTCAGTGTGTGCTTAAAGACAAATCTTGTCC 166
 75 AAGCCCGAGGTGTGACATAGAAAACCTGTTATGATCAACAACACCACTTCAA 134
 |||||||
 Db 167 AAGCCCGAGGTGTGACATAGAAAACCTGTTATGATCAACAACACCACTTCAA 226
 135 AATGTTACCGCAGCACAACGCCCAACAACACAGGGTGTGGAAGCCATTTCTGAAC 194
 |||||||
 Db 227 AATGTTACCGCAGCACAACGCCCAACAACAGGGTGTGGAAGCCATTTCTGAAC 286
 195 GTCCAGGATATGATCTCTGTAACACCTTTACTTAAATAAATGCAAGAGATTCTA 254
 |||||||
 Db 287 GTCCAGGATATGATCTCTGTAACACCTTTACTTAAATAAATGCAAGAGATTCTA 346
 255 CAACCTTAAGAAACAGAAACAAAGAAAGAACCGGCACTTTGGAGCGGAATGACAC 314
 |||||||
 Db 347 CAACCTTAAGAAACAGAAACAAAGAAAGAACCGGCACTTTGGAGCGGAATGACAC 406
 315 ACTGGCCCAAGGGCTACACCTGCTGTTCTTAAGAAATCTGGAATGTGGTCTCTTG 374
 |||||||
 Db 407 ACTGGCCCAAGGGCTACACCTGCTGTTCTTAAGAAATCTGGAATGTGGTCTCTTG 466
 375 GTGCCCTTACGAGTCTCTGATTAAACCAAGAGGAGTTCTCTCCGAGAGAAAGTGA 434
 |||||||
 Db 467 GTGCCCTTACGAGTCTCTGATTAAACCAAGAGGAGTTCTCTCCGAGAGAAAGTGA 526
 435 GCTAGCTTACGAGAGAAAGAGATGTGGTGTGAGAGGAAAGACTTTCAGCAGACGTG 494
 |||||||
 Db 527 GCTAGCTTACGAGAGAGAAAGAGATGTGGTGTGAGAGGAAAGACTTTCAGCAGACGTG 586

RESULT 4
 LOCUS BG866610
 DEFINITION 602785586F1 NCL_CGAP_SG2 Mus musculus cDNA clone IMAGE:4911884 5', mRNA sequence.
 ACCESSION BG866610
 VERSION BG866610.1 GI:14217150
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLAM10814 row: k column: 21
 High quality sequence stop: 718.
 Location/Qualifiers
 1..757
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="4911884"
 /clone_lib="NCL_CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-Sport6; Site: 1; Motif: Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCL_CGAP Library."
 BASE COUNT 236 a 193 c 164 g 164 t
 ORIGIN
 Query Match 97.1%; Score 516.4; DB 10; Length 757;
 Best Local Similarity 99.8%; Pred. No. 3e-105;
 Matches 517; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 15 ACACCCCCCGCAGAGATGTCAGTGTGTGCTTAAAGACAAATCTTGTCC 74
 |||||||
 Db 52 ACACCCCCCGCAGAGATGTCAGTGTGTGCTTAAAGACAAATCTTGTCC 111
 75 AAGCCCGAGGTGTGACATAGAAAACCTGTTATGATCAACAACACCACTTCAA 134
 |||||||
 Db 112 AAGCCCGAGGTGTGACATAGAAAACCTGTTATGATCAACAACACCACTTCAA 171
 135 AATGTTACCGCAGCACAACGCCCAACAACAGGGTGTGGAAGCCATTTCTGAAC 194
 |||||||
 Db 172 AATGTTACCGCAGCACAACGCCCAACAACAGGGTGTGGAAGCCATTTCTGAAC 231
 195 GTCCAGGATATGATCTCTGTAACACCTTTACTTAAATAAATGCAAGAGATTCTA 254
 |||||||
 Db 232 GTCCAGGATATGATCTCTGTAACACCTTTACTTAAATAAATGCAAGAGATTCTA 291
 255 CAACCTTAAGAAACAGAAACAAAGAAAGAACCGGCACTTTGGAGCGGAATGACAC 314
 |||||||
 Db 292 CAACCTTAAGAAACAGAAACAAAGAAAGAACCGGCACTTTGGAGCGGAATGACAC 351
 315 ACTGGCCCAAGGGCTACACCTGCTGTTCTTAAGAAATCTGGAATGTGGTCTCTTG 374

```

|||||
Db 352 ACTGCCCCAGGCTACCCCTGCCCCCTTGTCTAAGAAATCCCGAATGTGGTCCCTCC 411
Oy 375 CTGCTTCAGCAGCTCTGTGATTAACCAAGAGGAGTTCTCTCCGAGAGAAGTGA 434
Db 412 CTGCTTCAGCAGCTCTGTGATTAACCAAGAGGAGTTCTCTCCGAGAGAAGTGA 471
Oy 435 GCTACGTTAGCCAGAGAAGAAAGGATGTGGTGAGAGGAAAGACTTCGAGGAGCTG 494
Db 472 GCTACGTTAGCCAGAGAAGAAAGGATGTGGTGAGAGGAAAGACTTCGAGGAGCTG 531
Oy 495 CAAGAATCAAGCCTTTCATTAATTCCTGTCATCAT 532
Db 532 CAAGAATCAAGCCTTTCATTAATTCCTGTCATCAT 569

RESULT 5
LOCUS BC018582 1361 bp mRNA linear HTC 03-DEC-2001
DEFINITION Mus musculus, clone IMAGE:4163477, mRNA.
ACCESSION BC018582
VERSION BC018582.1 GI:11235764
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1361)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalobedcm.tmc.edu.
Villalob, D.K., Luna, R.A., Hale, S.M., Huliyk, S., Lu, X., Garcia,
A.M., Hollway, M., Teliford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 30 Row: m Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein
This clone has the following problem: no 5' EST match.
Location/Qualifiers
1. 1361
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4163477"
/tissue_type="Liver, normal, 5 month old male mouse."
/clone_lib="NCI_CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 401 a 330 c 328 g 302 t
ORIGIN
Query Match 97.1%; Score 516.4; DB 11; Length 1361;
Best Local Similarity 99.8%; Pred. No. 3.1e-105;
Matches 517; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 15 ACACCCCTGACAGAGATGTCCAGTGTGTTGCTTAGAGACAAAGTCACTTGTCC 74

```

```

|||||
Db 661 ACACCCCTGACAGAGATGTCCAGTGTGTTGCTTAGAGACAAAGTCACTTGTCC 720
Oy 75 AAGGCCAGGTGTGACCATAGAAACCTGTATGATTCACAAACAACACCATTTCAA 134
Db 721 AAGGCCAGGTGTGACCATAGAAACCTGTATGATTCACAAACAACACCATTTCAA 780
Oy 135 AATGTTACCGCAGACACAGACCCCAACACAGGATGTGGAGCAATTTCTGAAC 194
Db 781 AATGTTACCGCAGACACAGACCCCAACACAGGATGTGGAGCAATTTCTGAAC 840
Oy 195 GTCCAGATATGTACTCTTGTGTAACCACTTACCTTAACCAACATGCAAGAGATTCTA 254
Db 841 GTCCAGATATGTACTCTTGTGTAACCACTTACCTTAACCAACATGCAAGAGATTCTA 900
Oy 255 CAACCTAAGAAACAGAAAGAAACAAAGAACCCGACCTTTGGAGCCGGAATGACAC 314
Db 901 CAACCTAAGAAACAGAAAGAAACAAAGAACCCGACCTTTGGAGCCGGAATGACAC 960
Oy 315 ACTGGCCCCAGGGCTCAACCCCTGCTTGTCTAAGAAATCTGGAATGTGGTCTCTTG 374
Db 961 ACTGGCCCCAGGGCTCAACCCCTGCTTGTCTAAGAAATCTGGAATGTGGTCTCTTG 1020
Oy 375 CTGCTTCAGCAGCTCTGTGATTAACCAAGAGGAGTTCTCTCCGAGAGAAGTGA 434
Db 1021 CTGCTTCAGCAGCTCTGTGATTAACCAAGAGGAGTTCTCTCCGAGAGAAGTGA 1080
Oy 435 GCTACGTTAGCCAGAGAAGAAAGGATGTGGTGAGAGGAAAGACTTCGAGGAGCTG 494
Db 1081 GCTACGTTAGCCAGAGAAGAAAGGATGTGGTGAGAGGAAAGACTTCGAGGAGCTG 1140
Oy 495 CAAGAATCAAGCCTTTCATTAATTCCTGTCATCAT 532
Db 1141 CAAGAATCAAGCCTTTCATTAATTCCTGTCATCAT 1178

RESULT 6
LOCUS BG974340 874 bp mRNA linear EST 12-JUN-2001
DEFINITION 60284409P1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4979855,
cDNA sequence.
ACCESSION BG974340
VERSION BG974340.1 GI:14361977
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 874)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10978 row: k column: 24
High quality sequence stop: 846.
Location/Qualifiers
1. 874
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4979855"
/clone_lib="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; NotI;

```

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2.5 kb. Library constructed by Life
Technologies, catalog # 12018-016. Investigators providing
samples: Lohar Hennighausen/Priscilla Firth, NIH
Reference for transgenic model: Li et al., Cell Growth and
Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP
Library."

BASE COUNT 277 a 216 c 193 g 188 t

Query Match 96.8%; Score 514.8; DB 10; Length 874;
Best Local Similarity 99.6%; Pred. No. 6.9e-105;
Matches 516; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

15 ACACCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAGTCACTTGTGC 74
DB ACACCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAGTCACTTGTGC 259
75 AAGGCCAGGTGTGATCATTAGAAAACCTGTGTATGATCAACACACCATTC 134
DB AAGGCCAGGTGTGATCATTAGAAAACCTGTGTATGATCAACACACCATTC 319
135 AATGTTACCGCAGCAGACAGCCCAACACAGGAGGTGGGAAGCAATTTCTG 194
DB AATGTTACCGCAGCAGACAGCCCAACACAGGAGGTGGGAAGCAATTTCTG 379
195 GTCCAGATATGTACTCTTCTTACCACTTAAACATGCAAGAGATTTCTA 254
DB GTCCAGATATGTACTCTTCTTACCACTTAAACATGCAAGAGATTTCTA 439
380 GTCCGATATGTACTCTTCTTACCACTTAAACATGCAAGAGATTTCTA 439
255 CAACCTAGAAACAGAAAACAAAGAACCGGCACTTTGGAGCCGAAATGAC 314
DB CAACCTAGAAACAGAAAACAAAGAACCGGCACTTTGGAGCCGAAATGAC 499
440 CAACCTAGAAACAGAAAACAAAGAACCGGCACTTTGGAGCCGAAATGAC 499
315 ACTGCCCCAGGCTCAACCTGCTGTAGGAATCTGGAATGTGGTCTCTTG 374
DB ACTGCCCCAGGCTCAACCTGCTGTAGGAATCTGGAATGTGGTCTCTTG 559
500 ACTGCCCCAGGCTCAACCTGCTGTAGGAATCTGGAATGTGGTCTCTTG 559
375 CTGCTTACGAGAGTCTGTATTAACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 434
DB CTGCTTACGAGAGTCTGTATTAACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 619
560 CTGCTTACGAGAGTCTGTATTAACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 619
435 GCTACGTAGCCAGAGAGAGAGATGGGTAGAGAGAAAGACTTGGAGAGCTG 494
DB GCTACGTAGCCAGAGAGAGAGATGGGTAGAGAGAAAGACTTGGAGAGCTG 620
620 GCTACGTAGCCAGAGAGAGAGATGGGTAGAGAGAAAGACTTGGAGAGCTG 620
495 CAAAGAAATCAAGCCTTTCATTAATTCCTGTCATCAT 532
DB CAAAGAAATCAAGCCTTTCATTAATTCCTGTCATCAT 717

RESULT 7 841 bp mRNA linear EST 29-MAY-2001
LOCUS BG867642
DEFINITION 602787429F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913382 5',
mRNA sequence.
ACCESSION BG867642
VERSION BG867642.1 GI:14218182
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 841)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0818 row: j column: 07
High quality sequence stop: 801.
Location/Qualifiers

1. 841
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4913382"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (r1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-Sport6; Site_1:
NOT; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 266 a 215 c 178 g 181 t 1 others

Query Match 96.8%; Score 513.8; DB 10; Length 841;
Best Local Similarity 99.4%; Pred. No. 1.2e-104;
Matches 515; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

15 ACACCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAGTCACTTGTGC 74
DB ACACCCCTGCACAGAGATGTCAGTGTGCTTAGAGACAAGTCACTTGTGC 205
75 AAGGCCAGGTGTGATCATTAGAAAACCTGTGTATGATCAACACACCATTC 134
DB AAGGCCAGGTGTGATCATTAGAAAACCTGTGTATGATCAACACACCATTC 265
206 AAGGCCAGGTGTGATCATTAGAAAACCTGTGTATGATCAACACACCATTC 265
135 AATGTTACCGCAGCAGACAGCCCAACACAGGAGGTGGGAAGCAATTTCTG 194
DB AATGTTACCGCAGCAGACAGCCCAACACAGGAGGTGGGAAGCAATTTCTG 325
266 AATGTTACCGCAGCAGACAGCCCAACACAGGAGGTGGGAAGCAATTTCTG 325
195 GTCCAGATATGTACTCTTCTTACCACTTAAACATGCAAGAGATTTCTA 254
DB GTCCAGATATGTACTCTTCTTACCACTTAAACATGCAAGAGATTTCTA 385
326 GTCCAGATATGTACTCTTCTTACCACTTAAACATGCAAGAGATTTCTA 385
255 CAACCTAGAAACAGAAAACAAAGAACCGGCACTTTGGAGCCGAAATGAC 314
DB CAACCTAGAAACAGAAAACAAAGAACCGGCACTTTGGAGCCGAAATGAC 445
386 CAACCTAGAAACAGAAAACAAAGAACCGGCACTTTGGAGCCGAAATGAC 445
315 ACTGCCCCAGGCTCAACCTGCTGTAGGAATCTGGAATGTGGTCTCTTG 374
DB ACTGCCCCAGGCTCAACCTGCTGTAGGAATCTGGAATGTGGTCTCTTG 505
446 ACTGCCCCAGGCTCAACCTGCTGTAGGAATCTGGAATGTGGTCTCTTG 505
375 CTGCTTACGAGAGTCTGTATTAACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 434
DB CTGCTTACGAGAGTCTGTATTAACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 565
506 CTGCTTACGAGAGTCTGTATTAACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 565
435 GCTACGTAGCCAGAGAGAGAGATGGGTAGAGAGAAAGACTTGGAGAGCTG 494
DB GCTACGTAGCCAGAGAGAGAGATGGGTAGAGAGAAAGACTTGGAGAGCTG 625
566 GCTACGTAGCCAGAGAGAGAGATGGGTAGAGAGAAAGACTTGGAGAGCTG 625
495 CAAAGAAATCAAGCCTTTCATTAATTCCTGTCATCAT 532
DB CAAAGAAATCAAGCCTTTCATTAATTCCTGTCATCAT 663

RESULT 8 806 bp mRNA linear EST 21-NOV-2000
LOCUS BF299645
DEFINITION 602029327F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4164463 5',
mRNA sequence.
ACCESSION BF299645
VERSION BF299645.1 GI:11246168
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 806)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LLM9450 row: e column: 08
 High quality sequence stop: 724.
 Location/Qualifiers

FEATURES
 source 1..806
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:416463"
 /clone_lib="NCI_CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-Sport6; Site:1;
 Note: Site:2: Salivary gland; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 249 a 218 c 177 g 162 t
 ORIGIN

Query Match 96.2%; Score 511.6; DB 10; Length 806;
 Best Local Similarity 99.2%; Pred. No. 3.6e-104;
 Matches 514; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

15 ACACCCCTCTGCACAGAGATGTCAGTGTGCTTGTGAGACAAAGCACTTGTCC 74
 100 ACACCCCTCTGCACAGAGATGTCAGTGTGCTTGTGAGACAAAGCACTTGTCC 159
 75 AAGGCCAGGTGTGACATAGAAACCTGTATGATCAACAACACACCATTCAA 134
 160 AAGGCCAGGTGTGACATAGAAACCTGTATGATCAACAACACACCATTCAA 219
 135 AATGTTACCGCAGACACAGACAGCCAGAGGTGTGGAAGCCAAATTTCTGAC 194
 220 AATGTTACCGCAGACACAGACAGCCAGAGGTGTGGAAGCCAAATTTCTGAC 279
 195 GTCCAGATATGTACTCTGTACACCTTACTATAAAACATGCAAGAGGATCTA 254
 280 GTCCAGATATGTACTCTGTACACCTTACTATAAAACATGCAAGAGGATCTA 339
 255 CAACCTAAGAAACAGAAAGAAACAAAGAAACCGGCACTTTGGAGCCGGAATGAC 314
 340 CAACCTAAGAAACAGAAAGAAACAAAGAAACCGGCACTTTGGAGCCGGAATGAC 399
 315 ACTGCCCCAGAGGCTACACCTGCTGTCTAAGAAATCTTGAATGTGGTCTCTTG 374
 400 ACTGCCCCAGAGGCTACACCTGCTGTCTAAGAAATCTTGAATGTGGTCTCTTG 459
 375 CTGCTTACGAGATCTGTGATTAACCCAGAGGAGTCTCTCCGAGAGAAAGTGA 434
 460 CTGCTTACGAGATCTGTGATTAACCCAGAGGAGTCTCTCCGAGAGAAAGTGA 519
 435 GCTACGTTAGCAGAGAAAGAGGATGTGGTGAAGAGGAAAGACTTGCAGCAGCTG 494
 520 GCTACGTTAGCAGAGAAAGAGGATGTGGTGAAGAGGAAAGACTTGCAGCAGCTG 579
 495 CAAGAAATCAAGCCTTCAATATTCCTGTCATCAT 532
 580 CACAGAAATCAAGCCTTCAATATTCCTGTCATCAT 617

RESULT 9
 BF300481

LOCUS BF300481 822 bp mRNA linear EST 21-NOV-2000
 DEFINITION 602031613F1 NCI_CGAP_SG2 Mus musculus CDNA clone IMAGE:416858 5',
 mRNA sequence.
 ACCESSION BF300481
 VERSION BF300481.1 GI:11246991
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 822)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LLM9456 row: i column: 03
 High quality sequence stop: 782.
 Location/Qualifiers

FEATURES
 source 1..822
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:416858"
 /clone_lib="NCI_CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-Sport6; Site:1;
 Note: Site:2: Salivary gland; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 262 a 204 c 180 g 176 t
 ORIGIN
 Query Match 95.0%; Score 505.4; DB 10; Length 822;
 Best Local Similarity 99.6%; Pred. No. 8.6e-103;
 Matches 517; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

15 ACACCCCTCTGCACAGAGATGTCAGTGTGCTTGTGAGACAAAGCACTTGTCC 74
 119 ACACCCCTCTGCACAGAGATGTCAGTGTGCTTGTGAGACAAAGCACTTGTCC 178
 75 AA-GGCCAGGTGTGACATAGAAACCTGTATGATCAACAACACACCATTC 133
 179 AAGGCCAGGTGTGACATAGAAACCTGTATGATCAACAACACACCATTC 238
 134 AATGTTACCGCAGACACAGCCAGAGGTGTGGAAGCCAAATTTCTGAA 193
 239 AATGTTACCGCAGACACAGCCAGAGGTGTGGAAGCCAAATTTCTGAA 298
 194 GTTCAGAGATATGTACTCTGTACACCTTACTATAAAACATGCAAGAGATCT 253
 299 GTTCAGAGATATGTACTCTGTACACCTTACTATAAAACATGCAAGAGATCT 358
 254 ACAACCTAAGAAACAGAAAGAAACAAAGAAACCGGCACTTTGGAGCCGGAATGACA 313
 359 ACAACCTAAGAAACAGAAAGAAACAAAGAAACCGGCACTTTGGAGCCGGAATGACA 418
 314 CACTGCCCCAGAGGCTACACCTGCTGTCTAAGAAATCTTGAATGTGGTCTCTCT 373
 419 CACTGCCCCAGAGGCTACACCTGCTGTCTAAGAAATCTTGAATGTGGTCTCTCT 478
 374 GTTCCTTACGAGATCTGTGATTAACCCAGAGGAGTCTCTCCGAGAGAAAGTGG 433
 479 GTTCCTTACGAGATCTGTGATTAACCCAGAGGAGTCTCTCCGAGAGAAAGTGG 538
 434 AGCTACGTTAGCAGAGAAAGAGGATGTGGTGAAGAGGAAAGACTTGCAGCAGCT 493


```

|||||
Db 539 ACCTAGCTTAGCCAGAGAAAGGATGGGTGAGAGAGAAAGACTTCGACGACGT 598
QY 494 GCAAAATCAAGCTTCATATTCCTGTCAATCAT 532
Db 599 GCAAGAAATCAAGCTTCATATTCCTGTCAATCAT 637

RESULT 10
Bg666897 740 bp mRNA linear EST 29-MAY-2001
LOCUS 602786139P1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912329 5',
DEFINITION mRNA sequence.
ACCESSION Bg666897
VERSION Bg666897.1 GI:14217437
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 740)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10815 row: n column: 10
High quality sequence stop: 728.
Location/Qualifiers
1..740
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4912329"
/clone_1db="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1;
NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 230 a 189 c 158 g 163 t
ORIGIN
Query Match 94.8%; Score 504.4; DB 10; Length 740;
Best Local Similarity 99.6%; Pred. No. 1.4e-102;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 15 ACACCCCCCGACAGAGATGTCAGTGGTGTGTTAGACAGAAAGTCACTTTGTC 74
Db 46 ACACCCCCCGACAGAGATGTCAGTGGTGTGTTAGACAGAAAGTCACTTTGTC 105
QY 75 AAGGCCAGGTGTGACCATAGAAACTCTGTTATGATCACAACAACACACCATTCAA 134
Db 106 AAGGCCAGGTGTGACCATAGAAACTCTGTTATGATCACAACAACACACCATTCAA 155
QY 135 AATGGTTACCGACACACAGCCCAACACACAGGTTGTAAGCCAAATTCCTGAAC 194
Db 166 AATGGTTACCGACACACAGCCCAACACACAGGTTGTAAGCCAAATTCCTGAAC 225
QY 195 GTCCAGATATGTAATCTTGTAACACCTTACCTAAAAAATGCGAAGAGATTCCTA 254
Db 226 GTCCAGATATGTAATCTTGTAACACCTTACCTAAAAAATGCGAAGAGATTCCTA 285
QY 255 CAACCTTAAGAAAGAAAGAAAGAAAGAACCGGCACTTTGGACCGGAATATGAC 314
Db 286 CAACCTTAAGAAAGAAAGAAAGAAAGAACCGGCACTTTGGACCGGAATATGAC 345

```

```

QY 315 ACTGCCCCCAGGCTCACCCCTGCTTGTCTAAGAAATCTGGATGTGGTCTCTTG 374
Db 346 ACTGCCCCCA-GGCTCACCCCTGCTTGTCTAAGAAATCTGGATGTGGTCTCTTG 404
QY 375 CTGCTTCAGGATCTCTGTATTACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 434
Db 405 CTGCTTCAGGATCTCTGTATTACCCAGAGGAGTTCCTCCGAGAGAAAGTGA 464
QY 435 GCTAGCTTAGCCAGAGAAAGGATGTGGTGAGAGGAGAAAGACTTCAGAGACGTG 494
Db 465 GCTAGCTTAGCCAGAGAAAGGATGTGGTGAGAGGAGAAAGACTTCAGAGACGTG 524
QY 495 CAAGAAATCAAGCTTCATATTCCTGTCAATCAT 532
Db 525 CAAGAAATCAAGCTTCATATTCCTGTCAATCAT 562

RESULT 11
Bf537367 849 bp mRNA linear EST 11-DEC-2000
LOCUS 602050004F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4189464 5',
DEFINITION mRNA sequence.
ACCESSION Bf537367
VERSION Bf537367.1 GI:11624735
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 849)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9515 row: g column: 01
High quality sequence stop: 828.
Location/Qualifiers
1..849
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4189464"
/clone_1db="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1;
NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 271 a 207 c 193 g 178 t
ORIGIN
Query Match 94.8%; Score 504.4; DB 10; Length 849;
Best Local Similarity 99.6%; Pred. No. 1.4e-102;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 15 ACACCCCCCGACAGAGATGTCAGTGGTGTGTTAGACAGAAAGTCACTTTGTC 74
Db 135 ACACCCCCCGACAGAGATGTCAGTGGTGTGTTAGACAGAAAGTCACTTTGTC 193
QY 75 AAGGCCAGGTGTGACCATAGAAACTCTGTTATGATCACAACAACACACCATTCAA 134
Db 194 AAGGCCAGGTGTGACCATAGAAACTCTGTTATGATCACAACAACACACCATTCAA 253
QY 135 AATGGTTACCGACACACAGCCCAACACACAGGTTGTAAGCCAAATTCCTGAAC 194

```

```

Db 254 AATGTTACCGCAGCACACAGCCCAACACACAGGAGTGAGACCAATTTCTGAAAC 313
Oy 195 GTCCAGATATGATCTTGTGTACCACTTTACTATAAAACATGAGAGAGATTTCTA 254
Db 314 GTCCGATATGATCTTGTGTACCACTTTACTATAAAACATGAGAGAGATTTCTA 373
Oy 255 CAACCTAAGAAACAGAAAGAAACAAAGAAACACCGGACCTTTGGAGCCGGAATGACAC 314
Db 374 CAACCTAAGAAACAGAAAGAAACAAAGAAACACCGGACCTTTGGAGCCGGAATGACAC 433
Oy 315 ACTGCCCCCAGGAGCTACCTTGCCTTGTCTAAGCAATCTGGAATGTGGCTCTCTTG 374
Db 434 ACTGCCCCCAGGAGCTACCTTGCCTTGTCTAAGCAATCTGGAATGTGGCTCTCTTG 493
Oy 375 CTGCTTCAGGAGCTCTGTATTAACCAAGGAGGAGTTCTCTCCGAGAGAAAGTGGCA 434
Db 494 CTGCTTCAGGAGCTCTGTATTAACCAAGGAGGAGTTCTCTCCGAGAGAAAGTGGCA 553
Oy 435 GCTACGTTAGCCAGAAAGAAAGGAGTGGTGAGAGGAAAGACTTGCACGAGCTG 494
Db 554 GCTACGTTAGCCAGAAAGAAAGGAGTGGTGAGAGGAAAGACTTGCACGAGCTG 613
Oy 495 CAAGAAATCAAGCCTTTCAATATTTCCCTGCAATCAT 532
Db 614 CAAGAAATCAAGCCTTTCAATATTTCCCTGCAATCAT 651

```

RESULT 12
BF535022 891 bp mRNA linear EST 11-DEC-2000
LOCUS 602050476F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:418986 5',

DEFINITION mRNA sequence.
ACCESSION BF535022
VERSION BF535022.1 GI:11622385
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 891)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

Plate: LLM9516 row: g column: 19
High quality sequence stop: 739.
Location/Qualifiers
1. 891
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="418986"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

BASE COUNT 273 a 218 c 223 g 177 t
ORIGIN

Query Match 94.8%; Score 504.4; DB 10; Length 891;
Best Local Similarity 99.6%; Pred. No. 1.4e-102;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

Oy 15 ACACCCCTGACAGAGATGTCAGTGTGTGTGCTTAAAGACAAAGTCACCTTTGCC 74
Db 94 ACACCCCTGACAGAGATGTCAGTGTGTGTGCTTAAAGACAAAGTCACCTTTGCC 153
Oy 75 AAGGCCAGAGTGTGACCATGAAAACTGTGTATGATCTCAACACACACCATTTCAA 134
Db 154 AAGGCCAGAGTGTGACCATGAAAACTGTGTATGATCTCAACACACACCATTTCAA 213
Oy 135 AATGTTACCGCAGCACACAGCCCAACACACAGGAGTGAGACCAATTTCTGAAAC 194
Db 214 AATGTTACCGCAGCACACAGCCCAACACACAGGAGTGAGACCAATTTCTGAAAC 273
Oy 195 GTCCAGATATGATCTTGTGTACCACTTTACTATAAAACATGAGAGATTTCTA 254
Db 274 GTCCGATATGATCTTGTGTACCACTTTACTATAAAACATGAGAGATTTCTA 333
Oy 255 CAACCTAAGAAACAGAAAGAAACAAAGAAACACCGGACCTTTGGAGCCGGAATGACAC 314
Db 334 CAACCTAAGAAACAGAAAGAAACAAAGAAACACCGGACCTTTGGAGCCGGAATGACAC 392
Oy 315 ACTGCCCCCAGGAGCTACCTTGCCTTGTCTAAGCAATCTGGAATGTGGCTCTCTTG 374
Db 393 ACTGCCCCCAGGAGCTACCTTGCCTTGTCTAAGCAATCTGGAATGTGGCTCTCTTG 452
Oy 375 CTGCTTCAGGAGCTCTGTATTAACCAAGGAGGAGTTCTCTCCGAGAGAAAGTGGCA 434
Db 453 CTGCTTCAGGAGCTCTGTATTAACCAAGGAGGAGTTCTCTCCGAGAGAAAGTGGCA 512
Oy 435 GCTACGTTAGCCAGAAAGAAAGGAGTGGTGAGAGGAAAGACTTGCACGAGCTG 494
Db 513 GCTACGTTAGCCAGAAAGAAAGGAGTGGTGAGAGGAAAGACTTGCACGAGCTG 572
Oy 495 CAAGAAATCAAGCCTTTCAATATTTCCCTGCAATCAT 532
Db 573 CAAGAAATCAAGCCTTTCAATATTTCCCTGCAATCAT 610

```

RESULT 13
BG873314 743 bp mRNA linear EST 29-MAY-2001
LOCUS 602794169F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4925297 5',

DEFINITION mRNA sequence.
ACCESSION BG873314
VERSION BG873314.1 GI:14223854
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 743)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLMK10849 row: j column: 18
High quality sequence stop: 681.
Location/Qualifiers
1. 743
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4925297"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"

FEATURES
source

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 226 a 187 c 172 g 158 t

ORIGIN

Query Match 94.7%; Score 503.8; DB 10; Length 743;
Best Local Similarity 99.4%; Pred. No. 2e-102;
Matches 516; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

15 ACACCCCTGCACAGAGATGTCAGTGGTTGCTTACAGACAAAGTCACTTGTCC 74
|||||
31 ACACCCCTGCACAGAGATGTCAGTGGTTGCTTACAGACAAAGTCACTTGTCC 90
75 AAGGCCAGGTGTGACCATATGAAACCTGTATGATCAACACCAACACCATTCAA 134
|||||
91 AAGGCCAGGTGTGACCATATGAAACCTGTATGATCAACACCAACACCATTCAA 150
135 AATGTTACCGCAGACACACACCCACACAGGGTGTGGAAGCCAAATTTCTGTAAC 194
|||||
151 AATGTTACCGCAGACACACACCCACACAGGGTGTGGAAGCCAAATTTCTGTAAC 210
195 GTCCAGATATGTCTCTGTTACACCTTACTATAAAACATGCAGAAAGAGATTCTA 234
|||||
211 GTCCAGATATGTCTCTGTTACACCTTACTATAAAACATGCAGAAAGAGATTCTA 270
255 CAACCTAAGAAACAGAGAAACAAAGAAAGACCGGCACTTTGGAGCCGGAATATGAC 314
|||||
271 CAACCTAAGAAACAGAGAAACAAAGAAAGACCGGCACTTTGGAGCCGGAATATGAC 330
315 ACTGAGCCCGCAGGCTCAACCTCCCTTGTCTAAGAAATCTGGAATGTGGTCTCTTG 374
331 ACTGAGCCCGCAGGCTCAACCTCCCTTGTCTAAGAAATCTGGAATGTGGTCTCTTG 390
375 CTGCTTACGAGGATCTCTGATTAACCCAGAGAGATCTTCCCGAGAGAAAGTGGGA 434
391 CTGCTTACGAGGATCTCTGATTAACCCAGAGAGATCTTCCCGAGAGAAAGTGGGA 450
435 GCTACGTTAGCCAGAGAGAA-AGGATGTGGTGAAGAGGAAAGACCTTCGAGGACGT 493
|||||
451 GCTACGTTAGCCAGAGAGAAAGGAGATGTGGTGAAGAGGAAAGACCTTCGAGGACGT 510
494 GCAAGAAATCAAGCCTTTCATTAATTCCTGTCATATCAT 532
|||||
511 GCAAGAAATCAAGCCTTTCATTAATTCCTGTCATATCAT 549

RESULT 14
BF300873 988 bp mRNA linear EST 21-NOV-2000
LOCUS 602028873F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4163994 5',
DEFINITION mRNA sequence.
ACCESSION BF300873
VERSION BF300873.1 GI:11247396
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 988)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM9449 row: a column: 19
High quality sequence stop: 743.
Location/Qualifiers
1. '988

FEATURES
source

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4163994"
/clone_1lb="NCI_CGAP_SG2"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 319 a 244 c 241 g 184 t

ORIGIN

Query Match 94.5%; Score 502.8; DB 10; Length 988;
Best Local Similarity 99.4%; Pred. No. 3.3e-102;
Matches 515; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

15 ACACCCCTGCACAGAGATGTCAGTGGTTGCTTACAGACAAAGTCACTTGTCC 74
|||||
119 ACACCCCTGCACAGAGATGTCAGTGGTTGCTTACAGACAAAGTCACTTGTCC 178
75 AAGGCCAGGTGTGACCATATGAAACCTGTATGATCAACACCAACACCATTCAA 134
|||||
179 AAGGCCAGGTGTGACCATATGAAACCTGTATGATCAACACCAACACCATTCAA 238
135 AATGTTACCGCAGACACACCCACACAGGGTGTGGAAGCCAAATTTCTGTAAC 194
|||||
239 AATGTTACCGCAGACACACCCACACAGGGTGTGGAAGCCAAATTTCTGTAAC 298
195 GTCCAGATATGTCTCTGTTACACCTTACTATAAAACATGCAGAAAGAGATTCTA 254
299 GTCCAGATATGTCTCTGTTACACCTTACTATAAAACATGCAGAAAGAGATTCTA 358
255 CAACCTAAGAAACAGAGAAACAAAGAAAGACCGGCACTTTGGAGCCGGAATATGAC 314
|||||
359 CAACCTAAGAAACAGAGAAACAAAGAAAGACCGGCACTTTGGAGCCGGAATATGAC 418
315 ACTGAGCCCGCAGGCTCAACCTCCCTTGTCTAAGAAATCTGGAATGTGGTCTCTTG 374
|||||
419 ACTGAGCCCGCAGGCTCAACCTCCCTTGTCTAAGAAATCTGGAATGTGGTCTCTTG 478
479 CTGCTTACGAGGATCTCTGATTAACCCAGAGAGATCTTCCCGAGAGAAAGTGGGA 538
375 CTGCTTACGAGGATCTCTGATTAACCCAGAGAGATCTTCCCGAGAGAAAGTGGGA 434
|||||
435 GCTACGTTAGCCAGAGAGAAAGGAGATGTGGTGAAGAGGAAAGACCTTCGAGGACGT 494
539 GCTACGTTAGCCAGAGAGAAAGGAGATGTGGTGAAGAGGAAAGACCTTCGAGGACGT 598
495 CAAGAAATCAAGCCTTTCATTAATTCCTGTCATATCAT 532
|||||
599 C-AGAAATCAAGCCTTTCATTAATTCCTGTCATATCAT 635

RESULT 15
BG871007 741 bp mRNA linear EST 29-MAY-2001
LOCUS 602792051F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4923253 5',
DEFINITION mRNA sequence.
ACCESSION BG871007
VERSION BG871007.1 GI:14221547
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 741)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9abs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
<http://image.liml.gov>
Plate: LIML0844 row: e column: 14
High quality sequence stop: 740.
Location/Qualifiers
1. 741

FEATURES
source

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4923253"
/clone_lib="NCI_CGAP_S62"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Salivary gland; Vector: PCMV-SPORE6; Site_1:
Not1; Site_2: Sal1; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 231 a 192 c 166 g 152 t
ORIGIN

Query Match 93.3%; Score 496.4; DB 10; Length 741;

Best local similarity 98.6%; Pred. No. 8.8e-101;
Matches 511; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 15 ACACCCCTCCGACAGAGATGCCAGTGGTTGTCTAGACAGCAAGTCCTTGTCC 74
DB 201 ACACCCCTCCGACAGAGATGCCAGTGGTTGTCTAGACAGCAAGTCCTTGTCC 260
QY 75 AAGGCCAGGTGTGACATGAAACCTCTGTATGATCACAACAACACCATTCOA 134
DB 261 AAGGCCAGGTGTGACATGAAACCTCTGTATGATCACAACAACACCATTCOA 320
QY 135 AATGTTACCGGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 194
DB 321 AATGTTACCGGACACAGCCCAACACAGGAGTGTGAAGCCATTTCTGAAC 380
QY 195 GTCCAGATATGTACTCTTGTACACCTTACTTAAAAAATGACAGAGATTCTA 254
DB 381 GTCCAGATATGTACTCTTGTACACCTTACTTAAAAAATGACAGAGATTCTA 440
QY 255 CAACCTAAGAAACAGAAACAAAGAAACGCGCACTTTGGAGCGGAATGACAC 314
DB 441 CAACCTAAGAAACAGAAACAAAGAAACGCGCACTTTGGAGCGGAATGACAC 500
QY 315 ACTGCCCCAGGGCTCACCTGCCCTTGTCTAAGGAATCTGGAATGTGGTCTCTTG 374
DB 501 ACTGCCCCA-GGCTCACCTGCCCTTGTCTAAGGAATCTGGAATGTGGTCTCTTG 559
QY 375 CTGCTTACGAGTCTCTGATTAACCAAGAGAGTCTCTCCGAGAGAAAGTGA 434
DB 560 CTGCTTACGAGTCTCTGATTAACCAAGAGAGTCTCTCCGAGAGAAAGTGA 619
QY 435 GGTAGTTAGCCAGAAAGAAAGGATGGGTGAGAGGAAAGACTTCAGCGAGCTG 494
DB 620 GGTAGTTAGCCAGAAAGAAAGGATGGGTGAGAGGAAAGACTTCAGCGAGCTG 679
QY 495 CAAGAAATCAAGCTTTCATTAATTCCTGTCATCAT 532
DB 680 CAAGAAATCAAGCTTTCATTAATTCCTGTCATCAT 717

Search completed: October 19, 2002, 06:53:28
Job time : 1643 secs